METHODS TO EVALUATE THE EFFECTIVENESS OF CERTAIN SURROGATE MEASURES TO ASSESS SAFETY OF OPPOSING LEFT-TURN INTERACTIONS

A Dissertation
Presented to
The Academic Faculty

By

Lakshmi N. Peesapati

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy in
Civil and Environmental Engineering

Georgia Institute of Technology

August, 2014

METHODS TO EVALUATE THE EFFECTIVENESS OF CERTAIN SURROGATE MEASURES TO ASSESS SAFETY OF OPPOSING LEFT-TURN INTERACTIONS

Approved by:

Dr. Michael P. Hunter, Advisor
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Michael O. Rodgers
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Jorge A. Laval
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Angshuman Guin
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Gregory M. Corso
School of Psychology
Georgia Institute of Technology

Date Approved: June 20, 2014

Dedicated to my parents

P.V.S.N. Acharyulu and Nagamani

ACKNOWLEDGEMENTS

I should confess that it has been a long eventful journey till this point in my academic career. As I look back at the moment when I had to make a choice between coming to Georgia Tech to pursue PhD, and continuing with a software job, I am glad that I have taken the right decision. The discovery of new friends and colleagues, and a renewed appreciation for the continued pursuit of knowledge is something I value greatly.

I would like to express my gratitude to all my thesis committee members for their guidance on this research work. First of all, I would like to thank my advisor Dr. Michael Hunter for his constant guidance, personal attention, suggestions, and endless encouragement during my graduate studies. I have learnt a lot from him and he has been an outstanding mentor. He was and will always be my role model. I also feel very fortunate to work with Dr. Michael O Rodgers, whose technical and statistical guidance is invaluable. He is an inspiration for my interest in statistics, and for me choosing it as a career path. I am grateful to Dr. Angshuman Guin who has been of great help during the first phase of my PhD research in developing the data collection setup, driving to the site innumerable times and making the whole process work. He has always been quite approachable and the go-to person for any kind of help. I thank Dr. Jorge Laval for giving me an opportunity to have an in-class teaching experience and for his time and effort in serving on my doctoral committee. I also thank Dr. Gregory M. Corso for agreeing to be on my committee.

I would like to express my special thanks to Rama Chilukuri and his wife Bhanu who have been my family and gave me a “home away from home”. My gratitude also goes to all my friends around for their true friendship and encouragement. To name a few, my thanks go to Dr. Sandeep Bomminayuni, Sai Kumar Pallikonda, Dr. Bhargav Errangi, Dr. Manu Menon, Rajashekar, Ranjeeth, Rajat, Karthik, Uday, and Deepak who are friends for life and who gave me memories I will always cherish. I would also like to thank my colleagues at Georgia Tech, Dr. Saroch Boonsiripant, Dr. Seungkook Wu, Dr. Hoe Kyoung Kim, Dr. Wonho Suh, Dr. Dwayne Henclewood, Dr. Danjue Chen, and Popa Pratyaksa to name a few who have always supported me and made my stay at Georgia Tech pleasantly.

Finally, I would like to thank my wonderful family for their unconditional endless love and continued support. My dearest wife, Lavanya deserves my special gratitude for the enormous support, encouragement, and motivation she gave me during many tough moments of my life. My parents have always encouraged me to pursue more knowledge. Their love and enormous belief in me, sometimes more than I have on myself always inspires me. I like to express my gratitude to my parents in-law for their confidence in my abilities and for backing me. My sister Satya has been a source of enormous strength to me who has always stood by me. Without all of you, I would never have accomplished what I have today.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AADT	Annual Average Daily Traffic
CDF	Cumulative Distribution Function
EDF	Empirical Distribution Function
FHWA	Federal Highway Safety Administration
GDOT	Georgia Department of Transportation
GIS	Geographical Information System
GLM	Generalized Linear Model
GEV	Generalized Extreme Value
HSIP	Highway Safety Improvement Program
MSE	Mean Squared Error
PET	Post Encroachment Time
PDF	Probability Distribution Function
RCLINK	Roadway Characteristics Link
RSS	Residual Sum of Squares
RTM	Regression To Mean
SAFETEA-LU	Safety, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users

SUMMARY

"Highway Safety" is generally measured in terms of either the number of crashes or the rate of crashes. "Surrogate" on the other hand is a measurable or observable *non-crash event* that can either be converted or calibrated to crash frequency (Tarko A., (2009)). The National Highway Traffic Safety Administration (NHTSA) publishes a summary of crash statistics every year in the form of Traffic Safety Facts. These statistics show that in 2012, there were more than 5.6 million traffic crashes in the United States which caused more than 33,000 fatalities, and 2.3 million injuries. This shows the need for continuously evaluating the safety of our transportation facilities.

Traditionally transportation safety analyses have relied on crash data but this method has certain limitations. First is the limitation with timeliness. Crashes are relatively non-frequent events, and hence it would require multiple years of crash data to evaluate safety with certain confidence. The crash database relies on input from multiple agencies ranging from police reports to insurance agencies. Therefore, there is also a time-lag between when an incident takes place and the time when the crash database is updated with it. Crash reports are also sometimes inaccurate or incomplete and these inaccuracies are transferred into crash databases as well. However, most importantly, safety evaluation using crash data is retrospective in nature. This means that one has to wait for accidents to happen in order to evaluate safety, which means further loss of life and property. This limitation of crash data begets the question if there is any other measure that can help in the quicker evaluation of safety without waiting for further loss of life and property. The

idea behind using surrogates for safety analyses emanates from this requirement. Generally, in everyday traffic, there are close-calls or critical events where two or more vehicles can be on a collision path, requiring either or both the drivers to perform a maneuver to avoid that collision. These types of events have the potential to end up in a crash, are more frequent than the crashes but a majority of them do not end up in a crash primarily because of the intervention of one or more drivers. Nevertheless, these events still show the risk involved in these vehicular interactions. The hypothesis behind using surrogates is that there could be a correlation between these critical events and number of crashes at a location. If that is true, surrogates have a potential applicability in terms of faster evaluation of traffic safety. However, the predictive capability of surrogate measures is an area of ongoing research. Previous studies have often resulted in inconsistent findings in the relationship between surrogates and crashes, one of the primary reasons being inconsistent definition of a conflict. This often leads to considering a mixture of both serious and non-serious events observed in traffic and trying to establish a relationship with crashes.

The current research attempts to address this challenge and aims to increase confidence in the use of surrogate measures for safety evaluation. This study evaluated the effectiveness of certain surrogate measures (acceleration-deceleration profile, intersection entering speed of through vehicles, and Post Encroachment Time (PET)) in assessing the safety of opposing left-turn interactions at 4-legged signalized intersections by collection of time resolved video from eighteen selected intersections throughout Georgia. The research has been divided into four phases. The first three phases are related to each other where the results from the prior phase drive the research agenda for the next phase. The first phase deals with developing a data collection methodology for profile-based and point-based surrogate measures. The acceleration-deceleration profile and speed profile are considered to be profile-based measures as it is measured over a stretch of road from the trajectory of a vehicle. The PET measure is a point based measure as it is a single value to quantify the potential risk involved in the interaction between two vehicles. The results and lessons learnt from the first phase directed the agenda for the second phase that delved deeper into the effectiveness of PET as a surrogate measure to assess safety of opposing left-turn interactions. This phase considered paired comparison of PET data collected at two pairs of intersections. The third phase expanded this evaluation by collecting data at 14 more intersections having varied crash frequencies. The last phase of the research developed an online survey to experiment if human safety experts can classify the study intersections into safety categories, and evaluate how their classification compares with done based on crash frequency and PET.

Overall, this research demonstrated that surrogate measures can be effective in safety evaluation, specifically demonstrating the use of PET as a surrogate for crashes between left-turning vehicles and opposing through vehicles. The analysis of data found that the selected surrogate threshold is critical to the effectiveness of any surrogate measure. For example, based on the PET data collected at sample intersections considered in this study, the required PET threshold was found to be as low as 1.5 second to identify high crash intersections, and 1 second to have the best predictive power, significantly lower than the commonly reported 3 second and higher thresholds. Non-parametric rank analysis methods, generalized extreme value modeling, and generalized linear modeling techniques were used to model PET with other intersection and traffic characteristics to demonstrate the degree to which these surrogates can be used to identify potential high-crash intersections without resorting to a crash history. Finally, the effectiveness of PET and its assistance to decision makers is also been demonstrated through an example that helped find errors in reported crash data.

CHAPTER 1: INTRODUCTION

1.1 BACKGROUND AND MOTIVATION

Road traffic crashes are one of the world's leading public health problems. According to the National Highway Traffic Safety Administration (NHTSA), in 2012, approximately 5.6 million motor vehicle crashes occurred in the United States resulting in more than 33,000 deaths as well as substantial economic losses. A study jointly conducted by Cambridge Systematics and American Automobile Association (AAA) quantified the annual cost of road crashes in the United States to be $164.2 billion (Meyer, (2008)). These Figures show the need for constantly evaluating and improving the safety of our transportation facilities. Recognizing the importance of highway safety, the Federal Highway Administration (FHWA) has designated safety as one of the principal components of transportation management. At the federal level, the Highway Safety Improvement Program established by *Safe Accountable Flexible Efficient Transportation Equity Act: a Legacy for Users* (SAFETEA-LU) is a core program which aims to make significant progress in reducing highway fatalities.

Improving highway safety requires the availability of metrics to assess areas in need of remedial action and to evaluate the success or failure of any actions taken. Traditionally, safety of any transportation facility has been evaluated using crash data. However, this method has several limitations in terms of accuracy and efficiency, as explained below.

These limitations of crash data are considered to be general knowledge in the area of traffic safety.

- Crashes are rare events and crash data over an extended time frame (typically in the order of three years) is required to attain a meaningful level of confidence in the safety estimates (Nicholson, 1985). Crash data are also subject to “regression-to-mean” bias. RTM is a statistical phenomenon that introduces a bias in the selection of locations, when such a selection is based on only crash numbers, as crash numbers because of inherent randomness, tend to exhibit extremely high or extremely low numbers that are far from mean values in any chosen period.
- There is typically a significant time lag between the occurrence of a crash and the details of the crash being recorded into an accessible database. These limitations typically make assessment of the effect of a countermeasure on the safety of a facility a lengthy process.
- Crash data is generally based on multiple sources such as local and state police reports, and claims submitted to insurance agencies which are often inaccurate or incomplete. Data elements such as cause of crash, conditions of the site of crash, etc. are often judged by the police officer at the location of crash based on the available evidence or inputs from the people involved in the crash. These may be subjective in nature and inaccurate. Another source of inaccuracy can be the process of entering data into the crash database from the corresponding police report.
- Crash data does not provide insight into the pre-crash process. In other words, it does not provide sufficient information on the behavior of the vehicles before they were involved in the crash to fully assess potential countermeasures.

- Perhaps the most serious drawback is that crash data are retrospective in nature. This means that the approach of evaluating safety using crashes actually "waits" for accidents and that means further loss of life and property.

Given these limitations, it is highly desirable to identify other methods that can be used to evaluate safety in a more effective and/or faster manner. The broad problem is to find out if there is any measure which is more frequent than crashes, that is observable and measurable in the traffic system, and has a relationship with crashes to aid in evaluation of safety more rapidly.

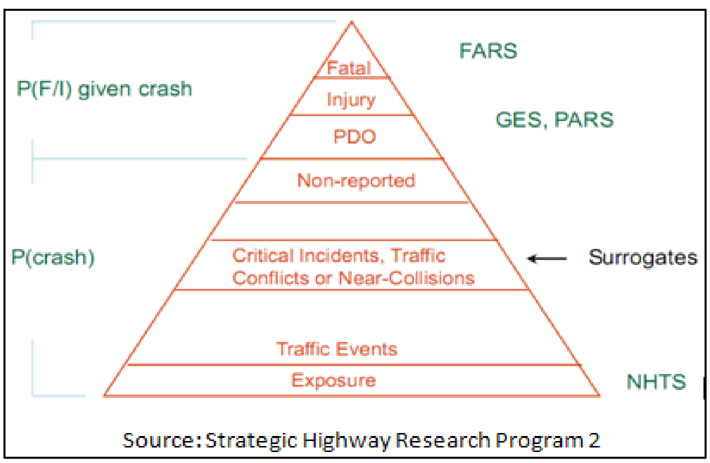


Figure 1.1: Traffic events severity and frequency (Campbell (2008)) (Not to scale)

Figure 1.1 shows the various safety related stages through which a vehicle may pass from entering the road to unfortunately getting involved in an incident. While a vehicle may not pass through all stages (for instance, injury and PDO are mutually exclusive) the pyramid represents increasing hazard, from general exposure to fatal incident. This pyramid was initially proposed by Hyden (1987). As soon as a vehicle enters the road, it is exposed to the vehicles around it. This vehicle interacts with the other vehicles on the road and these interactions are called traffic events. Some of these traffic events may turn into critical events where either or both of the vehicles needs to perform a maneuver to avoid a collision. An insufficient maneuver will lead to a collision. Depending on the severity of the crash, it may involve a fatality, injury, or only property damage. The significance of this triangle is that the area of each phase represents the frequency of events belonging to that phase. The Figure 1.1 is not to scale and it is just to show that the frequency of events decreases with increase in the severity of the event. For instance, fatal crashes are the most severe crashes and they are also the rarest, having the least area in the triangle. In this paradigm, the critical incidents or near-collisions which have a high potential of becoming a crash and are more frequent than the crashes, are considered to be surrogates for crashes.

Svensson (1998) proposed that though the pyramid proposed by Hyden (1987) is correct when all events on a road are considered, an extension of that pyramid is needed when only those events that are on a potential collision course are considered. Svensson proposed that, as in the case of speeds or accepted gaps which have tails on both sides of the distribution, even events on a collision course follow similar distribution with respect to severities. That is, if traffic interactions on a potential collision course are divided into three categories of severities: high, medium, and low, most of the interactions will be of medium severity while interactions having high and low severities will be very low (Svensson (2006)). That is, drivers generally take some risk in their decision making processes to gain mobility. Driver judgment in accepting gaps or a travel speed is essentially a trade-off between safety and efficiency, which makes the events having medium severity the highest frequency.

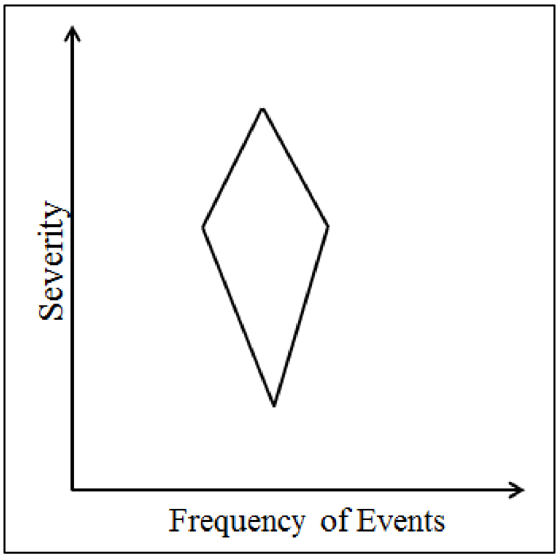


Figure 1.2: Extension of Hyden's pyramid for traffic interactions on a collision course (Svensson (2006))

This hypothesis alludes to a new direction for using traffic interactions for safety evaluation. Svensson hypothesized that in addition to the frequency of interactions having high severity, even interactions having low severity may be used as surrogates for safety evaluation (Figure 1.2). A crash is sometimes a result of an unexpected event that surprises the driver(s) involved. So, a site having high frequency of low severity events may also be considered negative as it does not give sufficient feedback about the risks involved. Hence the low caution that drivers may maintain at such locations might in fact turn an event of high severity, though rare, into a crash. In driver expectancy terms (the predisposition of people to believe that things will happen or be arranged in a certain way (Olson and Farber, 2003)) this becomes an event violation. That is, if an event never occurred before the driver fails to expect it now. So the overall shape of the distribution of the severity of events plays a major role in the evaluation of safety.

The use of surrogate safety measures allows for quicker safety analysis relative to the use of crash data. Surrogate measures are also considered in situations where historical crash data are limited or not available. Though there have been many surrogate measures of safety proposed in various studies (Gettman and Head, 2003), the definition still remains vague. According to the white paper by Tarko et al. (2009), surrogate measure of safety can be defined as "A measurable or observable non-crash event that is physically related in a predictable and reliable way to crashes and can be converted or calibrated into crash frequency and/or severity." Even Hyden's pyramid suggests that a surrogate measure is an "event" that is a near-crash or critical incident. However, surrogate measures generally used previously are actually "measures" to quantitatively assess the seriousness of the near-crash event. For example, application of sudden brakes to evade a potential crash could be the near crash "event", a way to measure its seriousness could be the "deceleration rate" with which the vehicle slows down. In this case, the near-crash event could be termed as a "surrogate event", and deceleration rate is the "surrogate measure" of safety.

The paper suggests that various factors that affect safety can be divided into two categories as shown in the Figure 1.3. The first category consists of the factors whose influence can be captured with a surrogate measure. For example, influence of insufficient sight distance or grade at an intersection can be measured using potential surrogate measures such as deceleration rate, acceleration noise etc. The second category consists of the other factors that affect safety such as driver expectancy, driver alertness, human factors, etc. whose influence may not be captured by a quantitative surrogate measure. Similarly, a mitigation measure applied at a facility may affect factors of safety belonging to either one or both of the categories.

Figure 1.3 below also depicts the relationship between surrogate measures and crashes. The horizontal arrow shows the causal relationship between surrogate measures of safety and crashes and the requirement that a surrogate event (critical event) occur for a corresponding crash to happen. The effect of a safety treatment on surrogate measures of safety is depicted by the vertical arrow. This means that if a treatment indeed affects a factor of surrogate measure, the treatment should also affect a surrogate event, where the affect is quantitatively assessed from the surrogate measure. The efficacy of a surrogate measure depends on how well it captures the relationship between a factor of safety and crashes. However, the requirement of causal relationship is still a matter of debate because there are a few studies that argue that surrogate measures can act as a "diagnostic" tool instead of a predictive tool which would then loosen the criteria for a surrogate measure that there should be a causal relationship and that frequency of a surrogate should be convertible to frequency of crashes.

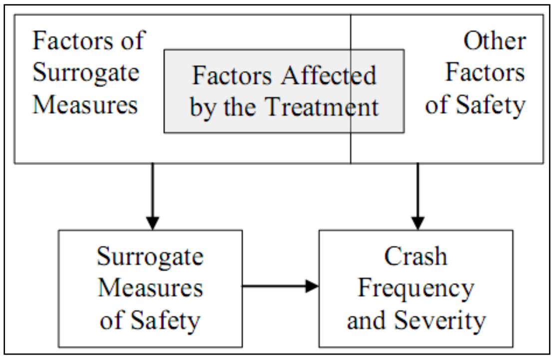


Figure 1.3: Relationship between crashes and surrogate measures of safety (Tarko et. al., (2009))

Changes in the number of observed conflicts (surrogate event) can give an early indication of the effectiveness of mitigation measures. The types of conflicts observed may also aid in selection of additional measures focused on addressing high conflict areas. So, conflicts allow for early estimate of the impact of a safety treatment and we may potentially evaluate the effect of a treatment on safety before accidents actually occur. Surrogate measures also may be used to aid in determining contributing factors to various crash types as they focus on pre-collision maneuvers and avoidance (Dingus et al., 2006). However, the important issue to be addressed here is how to identify the appropriate surrogate measure for a particular traffic interaction and demonstrate its linkage with crashes, thereby validating the effectiveness of the measure as a safety surrogate. There is a need to increase confidence, among practitioners and researchers, in using surrogate measures for faster evaluation of safety.

1.2 CHALLENGE

Though the literature and past research identified several measures which can potentially act as surrogates for safety, establishing the relationship between surrogate measures and safety is still a challenge. Previous studies have often resulted in inconsistent findings and models with low coefficients of determination (R^2 for the regression relationship between surrogates and accidents). The lack of statistically sufficient surrogate data, improper selection of study sites, consideration of inappropriate surrogate for the study interaction, errors in measurement, inconsistent definition of a conflict or critical event, and misapplication of statistical models may also have contributed to the current limited success of the evaluation of the effectiveness of surrogates.

The current research is aimed at addressing some of the above mentioned issues. This research focuses on three surrogate measures namely acceleration/deceleration values, intersection entering speed, and post encroachment time (PET). A review of literature shows that on-field data collection of profile-based surrogate measures has not been attempted much in the past, though with the recent advancements in video recording and computer vision techniques, there is growing interest in automated post-processing of videos to collect surrogate data. The aim of this research is to collect a statistically sufficient data sample for each of these surrogates, explore the collected data in more detail to understand their characteristics, and increase confidence on their broader applicability. This research also attempts to explore the use of expert knowledge in assessing the crash propensity at intersections. The interaction between left-turn vehicles and opposing through vehicles is being focused in this research, and the potential measures being studied will be evaluated for their effectiveness in acting as surrogates for this interaction.

The overall research has been divided into various phases and each of these phases follows a logical process of research. The study focuses on the crashes between left-turning vehicles and opposing through vehicle at signalized intersections. Phase 1 of the research developed a data collection methodology for both profile-based and point based surrogate measures, and used it to evaluate the effect of a treatment at a high speed multilane rural intersection. The experience and conclusions from phase 1 led to the idea behind phase 2 which looked at a potential surrogate called PET in more detail. Conclusions from phase 2 led to phase 3 which involved PET data collection at many more intersections and deeper analysis of that data to evaluate its diagnostic and predictive abilities. The last phase of research explored the potential of human expert evaluations in assessing traffic safety.

1.3 RESEARCH OBJECTIVES

- Develop a methodology to collect profile-based and point-based surrogate measures on-field. Understand the various issues in such methodology, and analyze various problems with the data. Evaluate the effectiveness of the developed methodology in an example case, where surrogate measures such as acceleration/deceleration values, post encroachment time, and intersection entering speed of through vehicles are used in the evaluation of safety treatment at an intersection in rural Georgia.

- Expand the evaluation the effectiveness of Post Encroachment Time as a surrogate measure of safety focusing on the interaction between a left-turning vehicle and opposing through vehicle, while understanding its limitations and applicability.
- Develop models in terms of PET data and other intersection operational and geometric characteristics to model crash data. Compare the potential of PET in terms of its diagnostic vs. predictive ability.
- Experiment if an expert panel can successfully classify intersections into “high”, “medium”, and “low” categories and evaluate the relative safety of intersections by manually synthesizing visual information and then rate the intersections according to potential for crashes. Explore and analyze the similarities and differences if any between the assessments of intersections provided by the expert panel, that shown by crash data, and PET.
- Identify needs for future research in this area.

1.4 REPORT ORGANIZATION

The main part of the dissertation is organized into eight chapters. These chapters (especially chapters 4, 5, 6, and 7) provide Figures only for one or two sample cases (intersections) for demonstration purposes. All other data and Figures corresponding to other intersections, any other background material relevant to the research and analysis is presented in the corresponding appendix mentioned in the chapters.

Chapter 1: Provides an introduction and perspective to the research, its importance and motivation on a broader scale, and specific objectives of the current study.

Chapter 2: Explores and summarizes the past research conducted in the area of surrogate measures for safety, the identified measures, data collection methods, and statistical models used to establish their relationship with crashes.

Chapter 3: Describes phase 1 of the project, the data that was collected, smoothing algorithm applied to reduce noise in the raw data, before after treatment analysis and how the results of this phase led to phase 2.

Chapter 4: Describes briefly the objective of phase 2, the method used to select locations, data collection process, and the relevant results that led to phase 3 of the study. This chapter also discusses phase 3 (basically an extension of phase 2), the locations selected for this study, the data collected, and the analysis approach used to accomplish the objectives of this research.

Chapter 5: "Analysis Chapter" that describes in detail about the various models used to establish the relationship between surrogate measure (PET), characteristics of the intersection, and crashes.

Chapter 6: Describes about the survey which was developed to exploit expert panel's assessment to evaluate safety. The chapter discusses the results obtained from the survey and how they relate to the crash data.

Chapter 7: Summarizes the conclusions drawn from the various phases of the project and corresponding analyses performed , overall contributions of the research, and future needs and directions.

CHAPTER 2: LITERATURE REVIEW

Highway safety evaluation has traditionally been performed using incident data, as an incident is a direct measure of lack of safety. However, it is understood and well accepted that evaluating safety through incidents has limitations in terms of timeliness and efficiency. First of all, crashes are rare events and any safety evaluation using crash data requires an extended time frame (typically in the order of three years (Nicholson, 1985)). As explained in chapter 1, regression-to-mean bias, and time-lag between the occurrence of an incident and its corresponding entry into an accessible database are other limitations with respect to timeliness. Quality is another issue with using crash data. Some errors occur while processing police reports into a more useable database format while others occur due to subjectivity involved in determining certain elements of crash data. However, the most important limitation of using crash data is that it is a retrospective approach which means that one has to wait for incident to happen in order to evaluate safety which means further loss of life and property.

Given these limitations, it is highly desirable to identify other methods that can be used to evaluate safety in a more effective and/or faster manner. The discussion in chapter 1 (especially the traffic-events pyramid) laid the basis for identifying measures that can be used for indirect evaluation of safety (crash data being the direct measure). The concept is that near-crash events can be used as surrogates for crashes as they indicate partial breakdown in traffic. These events are considered to be "surrogate measures of safety".

The term “surrogate measures for crash data” is also often used as these measures are used in place of crash data in situations where crash data is not available or for quicker evaluation of safety (Gettman et al., 2008). In addition to near-collisions, many studies suggest the use of traffic operational characteristics as surrogates for safety and more details follow in section 2.1.

With this preface, it follows that to accomplish the objectives of this research, a review of the following important topics is necessary:

- Knowledge of various surrogate measures identified in previous research efforts is the first requirement. It is important to understand the prior experiences of researchers with using surrogates and the conclusions made about their effectiveness, methodologies adopted to collect such data, strengths and limitations of these methodologies, etc.
- The traffic-events pyramid described in the previous chapter shows the potential for near-crashes to act as surrogates for crash data. It is necessary to understand what under what conditions a traffic event should be recognized as a near-crash or a serious conflict.
- Various statistical techniques used by previous researchers to evaluate safety of transportation systems by modeling crashes in terms of surrogate measures and factors of safety.
- Selection of candidate intersections such that the selection process does not bias the evaluation of surrogate measures.

Each of the above topics is reviewed in the following sections of this chapter.

2.1 SURROGATE MEASURES OF SAFETY

The use of surrogate safety measures is expected to allow for more rapid and earlier safety analysis relative to studies using actual crash data. This section lists various surrogate measures identified previously in various studies, discusses the corresponding data collection methodologies, and evaluates their advantages and limitations. Finally a summary of the important findings and conclusions from this section of literature review is presented.

Most of the surrogate measures proposed in the previous studies relate to traffic operational characteristics. These measures may be categorized as macroscopic or microscopic. Many of the macroscopic measures, particularly for intersections, include fairly standard measures of effectiveness. Some of these are delay, travel time, red light violations, stop-bar encroachments, queue length (Perkins and Bowman, 1986; Gettman and Head, 2003), traffic oscillations (Zheng et al., 2010), and parameters of a two-fluid model (Dixit et al., 2011). Gettman and Head (2003) stated that, “No attempt was made to relate these measures quantitatively to crash rates, but rather to assert some rules-of-thumb” about some of these macroscopic measures.

Much of the literature however identifies microscopic measures as surrogates. The basic intuition is that the use of frequency of conflicts or near-crashes, which are events involving individual or pairs of vehicles, are microscopic in nature. Previous studies suggest the use of braking or swerving (Perkins and Harris, 1967), Gap Time, Post Encroachment Time (Allen et al., 1978), Time-to-Collision (Hayward, 1972), etc. It can be seen that each of these are measured at a particular time or spot during the trajectory of a vehicle, for example, at the moment when the conflicting vehicle applies brakes. Hence these are called point measures. However, there are others microscopic measures that are considered over some amount of time or distance that a vehicle travels. Examples of such identified microscopic surrogate measures are speed variance (Lave, 1985), acceleration noise (Herman et al., 1959), Extended Time-to-Collision and Time Integrated Time-to-Collision (Minderhoud and Bovy 2001), etc. Since these measures deal with vehicle trajectories or profiles, they are called profile-based surrogate measures. Each of these measures has strengths and limitations and is collected using various methods. The following sub-sections delve deeper into some of these potential surrogates as the first important step in proceeding with the current research objectives.

2.1.1 Speed and its variations

One of the earliest factors considered to be affecting safety is speed. There are numerous studies that evaluated the relationship between speed and crashes (Haddon (1961), Solomon (1964), (Kockelman and Murray (2007), Boonsiripant et al. (2011)). According to Haddon (1961), the relationship between speed and safety can be divided into two terms, one related to "pre-event" phase and the other related to "event" phase. The pre-event phase examines the effect of speed on the probability of accident and the event phase deals with the accident severity. Since the pre-event factor concerns the likelihood of a crash, speed is a potential surrogate safety measure. One of the earliest attempts to evaluate the relationship between speed and crashes was undertaken by Solomon (1964) where he compared pre-crash traveling speeds with speed of vehicles under normal conditions. It was found that large speed variations in either direction of the average speed are positively related to highway crashes. Other studies (Kockelman and Murray, 2007; Kloeden et al., 1997) have also evaluated the potential of pre-crash speed as a surrogate for safety but had contrary findings.

Variations of speed measures such as 85th percentile speed, speed variance, Skewness Index etc. have also been studied. Studies such as Garber and Gadiraju (1989) and Lave (1985) concluded that the crash rates increased with increasing speed variance among drivers. Boonsiripant et al. (2011) explored the use of speed variation over a road segment (profile based measure) as a surrogate to crash frequency of a facility. The continuous speed profile was obtained from GPS-equipped vehicle data. The study found that the most important explanatory variable in the crash prediction models is acceleration noise. The measure "stop frequency" was also found to have a positive relationship with crash frequency. However, various other speed measures considered in this study such as mean and standard deviations of 85th percentile speed, and speed bands did not show significant relationship with crash frequency. However, the study itself mentions certain limitations with respect to selection of corridors and the need to incorporate more geometric variables in the crash prediction model.

It still remains a matter of debate whether speed can be considered as a surrogate measure for safety. According to the white paper on surrogate measures of safety (Tarko et al. (2009)), since crashes are measured in terms of a frequency, an effective surrogate would also be an event whose frequency would be a measure of safety. Speed as such is not a frequency measure and hence it may be difficult to convert the change in speed to change in crashes. In contrast, measures such as number of vehicles traveling above a certain speed, number of speeding tickets etc. are frequencies and can be considered as surrogates if sufficient correlation with crashes can be identified.

2.1.2 Braking and Swerving

Several of the earliest measures considered to be surrogates for safety are braking and swerving, evasive actions taken by the drivers to avoid a collision. Researchers over the years have tried to identify indicators which would give the frequency of conflicts thereby acting as surrogates for safety. These actions are assumed to be indicators of traffic conflicts where either or both of the involved drivers needed to perform these maneuvers to evade a potential crash. The methodology to collect such data is called Traffic Conflict Technique (TCT). TCT was initially developed in the 1960s and is still being used today with some modifications. Much of the literature available to date focuses on TCT as a surrogate safety measure (Perkins and Harris (1967), Sayed et al. (1999), Parker et al. (1989), Hyden (1987)).

An advantage of such an approach is that the conflicts can be detected in the field in real time and hence the data collection process takes a very short time. But the method requires manual recognition and judgment which is inherently a subjective process. When a driver performs an evasive action, it is difficult to determine if the action was precautionary or if there existed a conflict. Secondly, there is a potential that some conflicts may have been missed by the field observer. Thirdly, evasive action might be absent even in some close conflicts as this process also involves driver judgment. Finally, large scale studies, especially involving intersections require many human observers looking at various approaches to an intersection and hence are not cost-effective (Glauz and Migletz, 1980).

Recognizing the limitations of the TCT methodology, various research efforts have been undertaken to reduce subjectivity in the data collection process. Researchers have tried to quantitatively recognize conflicts in the place of qualitative and subjective identification. The surrogate measures explained below incorporate some improvements over traditional TCT.

2.1.3 Gap Time

Gap Time is the interval between expected arrival times of the conflicting vehicles at the area of conflict if they continued with same speed and path (Gettman et al. (2009)). It should be understood that it is the perceived time and not an actual observed time. For example, drivers of left-turning vehicles perceive an available gap time during permissive left-turn phase and take a decision either to accept or reject the gap. However, gap time in the context of surrogate measures is not what is perceived by the driver, but something that can be quantitatively computed based on the current location and speeds of the involved vehicles. This means that a gap time value of 0 can be computed for vehicles that are not involved in a crash at the end of their interaction if they were initially on a collision path at the time of computing the gap time but eventually performed evasive maneuvers. However, computing Gap Time is not straightforward. It requires computing the expected time of arrival of both the conflicting vehicles which requires maintaining location and speed data of both vehicles.

2.1.4 Encroachment Time, and Post Encroachment Time

Encroachment Time is the time duration during which the turning vehicle infringes upon the right-of-way of through vehicle (Allen et al., 1978). As the definition suggests, this measure is only concerned with the left-turning vehicle to determine the pace with which it makes the turn. However, Encroachment Time by itself does not signify any crash proximity or measure of safety. In case of large gaps available, large Encroachment Times do not pose any safety challenge. On the other hand, Post Encroachment Time (PET) is a measure that considers the arrival of the conflicting vehicle as well. PET is the time interval between end of encroachment of one vehicle and the time at which the conflicting vehicle actually arrives at the potential point of conflict (Cooper and Ferguson, 1976; Allen et al., 1978).

PET is one of the most popular surrogate measures because of the ease with which it can be measured. The first advantage of the measure PET is that it quantitatively shows the crash proximity of an interaction or conflict. By definition, a PET value of 0 implies a crash. The closer the PET value to 0, the less time that separated the events from being a crash. Moreover, PET data collection is also relatively straightforward. It requires two time stamps (one when the first vehicle leaves the area of encroachment and the second when a conflicting vehicle enters the area of encroachment) to compute PET. However, PET does not indicate whether the encroachment time was a result of a conflict in which the drivers accepted the small gap or if it was the result of evasive maneuver like braking or accelerating performed by either or both drivers (Chin and Quek, 1997). A variation of PET was also proposed by Allen et al. (1978) that takes into account the expected time of arrival of the opposing through vehicle at the area of conflict based on its current speed. This measure was termed "Initially attempted PET" and was defined as "the time lapse between commencement of encroachment by turning vehicle plus the *expected* time for the through vehicle to reach the point of collision and the completion time of encroachment by turning vehicle" (Gettman et al. (2009)). A PET value on the other hand is the final value observed that might have been the result of an evasive maneuver applied by the driver to avoid a collision. Both Allen et al. (1978) and Cooper and Ferguson (1976) have concluded that PET has the best correlation with crashes.

2.1.5 Proportion of Stopping Distance

Besides the time-based measures, some other measures that explain spatial or kinematic characteristics of traffic interactions were proposed in previous studies (Debnath and Chin, 2010). One such measure is the Proportion of stopping distance (PSD) which is defined as the ratio of distance available to maneuver to the distance remaining to the projected location of collision (Gettman and Head, 2003). This measure was first found to be mentioned and measured by Allen et al. (1978). By definition, a minimum PSD value of 1 is required to stop safely. This measure however has not been explored enough in previous studies.

2.1.6 Acceleration Noise

The term “Acceleration Noise” was first coined in a paper by Herman et al. (1959) which investigated the propagation of disturbances down a line of vehicles on a highway. Acceleration noise was first proposed as a parameter that can be used to characterize various drivers and road conditions. The authors also hinted at the possible relationship between this parameter and safety hazardousness of the situation. However, the hazard might be due to road conditions or the individual driver’s behavior. A 1958 Chevrolet equipped with a Statham accelerometer was used to measure the acceleration noise under various driving conditions. Though the experiment showed that acceleration noise can be used to identify aggressive drivers, it does not show any evidence of its applicability as a surrogate measure for safety.

Jones and Potts (1962) conducted the first study completely focusing on acceleration noise. According to their paper, acceleration noise, denoted by σ, is defined as the root mean square of the acceleration, so that

$$\sigma^2 = \frac{1}{T}\int_0^T (\mathrm{a(t)} - \mathrm{a_{av}})^2 \, dt \text{ , and} \tag{2.1}$$

$$a_{av} = \frac{1}{T}\int_0^T a(t)dt = \frac{1}{T}(v(T) - v(0)) \tag{2.2}$$

where *v(t)* and *a(t)* are the speed and acceleration of a car at time t and a_{av} is the average acceleration of the car for a trip time T. To quicken the process of data collection, a "Tachograph" that plots speed time graph was used on the test car. The results from this study were used to determine how road conditions and traffic conditions affect acceleration noise. Though this study was not aimed at finding the relationship between acceleration noise and safety, the authors concluded that high values of acceleration noise indicate a potentially dangerous situation and take the example of one of their study locations to illustrate some correlation between acceleration noise and accident potential.

Sattari and Powell (1987) considered acceleration noise and the mean velocity gradient as safety indicators. In contrast to the previous studies, this study considered acceleration noise with respect to spatial increments instead of time increments. This study showed that both acceleration noise and mean velocity gradient offer useful correlations with accident risk, and that the correlation coefficients improve when averaging of data is based on larger section lengths. They used an onboard servo-accelerometer mounted horizontally and aligned in the longitudinal direction of the motion to collect field data. Though the results suggest the applicability of these measures as surrogates for safety, there were only eight data points owing to the difficulty in measurement.

2.1.7 Deceleration Rate

As shown in the previous section 2.1.6, deceleration rate has been studied previously under acceleration noise (which includes both accelerations and decelerations). There are two other definitions of deceleration rate in surrogate safety literature. Allen et al. (1978) defined it as the maximum deceleration applied by the conflicting driver which is a measure of his perceived hazard ahead.

Chin et al. (1992) however defined deceleration rate as an average deceleration that a conflicting or crossing vehicle is required to take to avoid a collision which they term as "deceleration-to-avoid-collision" in case the vehicle in the front does not change its course or speed. They used this definition in the case of merging traffic at an expressway. The distribution of deceleration-to-avoid-collision values was combined with skid resistance values to find the probability of a serious conflict. Though they found that the probability can vary anywhere between 7 in 100,000 to 6 in 10,000 for different periods of observation, no study of relationship with crash data was done to validate the measure. The paper however concludes that the data collection process is labor intensive and automation is required to make it efficient.

2.1.8 Time-To-Collision and its variants

The concept of time-to-collision (TTC) was first introduced by Hayward in 1972 when he coined the term "time measured until collision". According to his paper (Hayward, 1972), the aim of this term was to introduce objectivity into the identification of conflicts by applying a scale of danger to near-miss situations. According to Hyden (1996), a TTC value at an instant t is defined as the time that remains until a collision between two vehicles would have occurred if the collision course and speed difference are maintained. See Figure 2.1 for illustration of time-to-collision.

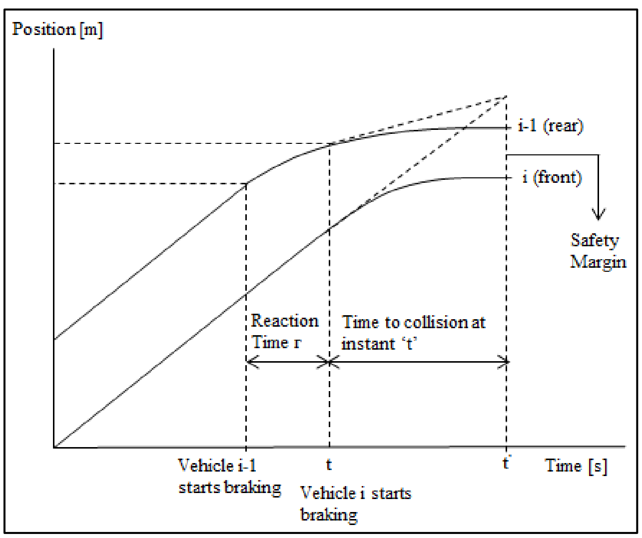


Figure 2.1: Illustration of time-to-collision between two vehicles (Minderhoud and Bovy, 2001)

In the above depicted situation, the lead vehicle applies brakes first and then it takes a reaction time τ for the following vehicle to react and apply brakes. TTC can be calculated the moment the speed of the following vehicle is greater than that of the lead vehicle. At that instant of time, the trajectories of both vehicles are projected assuming the same speeds to exist and TTC is the time it takes for the projected trajectories to meet which implies a collision. In the above illustration, the time-to-collision at time t is shown.

It can be inferred from the definition of TTC that lower the value of TTC, greater the proximity to crash and hence greater the severity of conflict. Also it can be understood that the lowest TTC is observed just before the instant the following vehicle applies its brakes. With this idea, Hyden (1976) conducted a study where he used trained personnel to count serious conflicts by observing tape recordings collected at two different set of intersections. Though both of these sets showed good correspondence of the ratio (serious conflicts/accidents), there is still subjectivity involved in the identification of serious conflicts. Moreover, the number of years of accident data considered is prone to regression-to-mean bias. A similar study was conducted by Cooper (1977) at merging locations on expressways. However, this study concluded that there is no relationship between serious conflicts identified by low TTC values and accident history.

Modifications to the original TTC measure have been suggested in the paper Minderhoud and Bovy (2001). In contrast to the classical TTC values that are measured at a specific instant of time which is application of brakes by the following vehicle, the new indicators use vehicle trajectories collected over a specific time period to calculate a general safety indicator value. The first of these new indicators is the Time Exposed TTC (TET) which measures the length of time that a vehicle involved in a conflict spends under a threshold

TTC value during a specified time period. However, it can be seen that the TET measure does not consider how low the observed TTC values are below the threshold. This limitation is removed in the second indicator which is Time Integrated TTC (TIT) which uses the integral of the TTC profile of drivers to express the actual “level of unsafety” over the specified time period. However, TTC is inherently difficult to measure because it requires an observation of when the following vehicle applies the brake, and information about speeds of both vehicles at that instant of time. Computation of measures such as TET and TIT is more difficult because it requires saving vehicle trajectories. Processing recorded videos to obtain trajectories is a very time consuming and labor intensive process. Until there are reliable automated vehicle tracking systems, simulation is the only way such measures can be computed, but simulation has its own limitations. Moreover, the effectiveness of these measures has not been consistently validated with crash data.

TTC distributions have also been applied in several other studies to identify traffic safety impacts (Hogema and Janssen, 1996; Fancher et.al., 1997; Van Arem & De Vos 1997; Vogel, 2002; Kiefer et al., 2004).

2.1.9 Other Studies

The potential to derive various surrogate measures of safety from existing microscopic simulation models has also been investigated in a FHWA study (Gettman and Head (2003)). Various algorithms were developed to compute surrogate measures and detect conflicts. Surrogate Safety Assessment Model (SSAM) was developed to detect conflicts automatically from vehicle trajectories generated from simulation models such as those from VISSIM, PARAMICS, AIMSUM and TEXAS. A field validation test was performed by considering 83 intersections (all four-leg, urban, signalized intersections) by modeling them in VISSIM, collecting conflict data from SSAM and then by comparing the output from SSAM with actual crash histories (Gettman et al., 2008). It was found that the conflicts data obtained were correlated with the crash data collected in the field, with the exception of conflicts during path-crossing maneuvers, which they observe, are under-represented in the simulation. The relationship between total conflicts and total crashes exhibited a correlation (R-squared) value of 0.41. Moreover, a good correlation between conflicts and crashes was not found for intersections having high number of conflicts and crashes. The study concluded that though the SSAM showed potential, the validation effort did not show conclusive results. After the introduction of SSAM, there have been many studies trying to validate the model and to compare the effectiveness of simulation vis-à-vis field data collection.

In the last decade, “naturalistic driving studies” have caught the attention of safety researchers. Shankar et al. (2008) define these studies as *those undertaken by using unobtrusive observation or observation taking place in a natural setting*. Driver actions under various conditions, especially before a crash are captured using vehicle-based sensors and video instrumentation. The “100-Car Naturalistic Driving Study” is the first instrumented vehicle study undertaken in 2001 with the aim of collecting such data on a large scale (Dingus et al., 2006). However, at the end of the study period, very few crashes were observed. Recently in 2010, under the SAFETEA-LU program, a bigger study called "SHRP 2 Naturalistic Driving Study" was authorized and is managed by Transportation Research Board (TRB) on behalf of the National Research Council (NRC) (www.shrp2.nds.us). These studies are especially encouraging for collecting surrogate data as they enable the availability of real data before a crash. The 100-Car Naturalistic Driving Study gave very valuable information regarding the relationship between crashes and near-crashes (Guo et al., 2010) as follows:

- No significant differences in the number of contributing factors for both crashes and near crashes and that there is a positive relationship between crashes and near-crashes.
- The use of near-crashes as surrogates can significantly improve the precision of risk estimation, though the estimate would be conservative.

Recent advancements in video and image processing technology have paved the way for automated detection of vehicles, tracking their motion, and determining various traffic parameters (Chin et al., 1992; Beymer et al., 1997; Vasquez et al., 2004; Kanhere et al., 2006; NGSIM, 2008). Automated post processing of a video recording helps reduce the variation in results due to human interpretation and subjectivity. A review of the literature shows that computer vision technology has also been used in detection of conflicts, calculating speeds of vehicles, and data collection for a few surrogate measures like Post Encroachment Time (PET) and Gap Time (Saunier and Sayed, 2007; Saunier and Sayed, 2008; Ismail et al., 2010; Laureshyn et al., 2010; Autey et al., 2012) and there is growing interest and potential in this approach. However, there still are various additional obstacles associated with a completely automated detection methodology. These include occlusion, vehicle and headlight reflections, false calls etc. Though recent research has tried to overcome many of these problems, limitations still exist in terms of traffic conditions, camera interface requirements and associated equipment investments. For example, one of the recent studies that successfully implemented such a system (Aubin et al., 2013) has limitations in the case of dense and turbulent traffic flows, and hence their methodology targets high-speed, low to medium-flow scenarios only. Therefore such methods need to be completely validated under various conditions to be put into practice by wider researchers and practitioners.

2.1.10 Section Summary

Surrogate measures are expected to provide an alternative and timely method of safety evaluation. There is a considerable amount of work done in identification and collection of surrogate data but these studies at best provide mixed conclusions about their effectiveness and applicability. There are three major issues with their application. First of all, there is still a debate going on in research circles about the definition of a surrogate and criteria for recognizing one. For example, there are differing views about speed being considered as a surrogate measure for safety. The second concern is with respect to difficulties in collecting data. The TCT method that was proposed in the 1960s has subjectivity inherent in the process and hence later studies tried to reduce subjectivity by identifying measures that enable quantitative identification of near-crashes. Point based measures such as TTC, PET, deceleration rate, and gap time are relatively easier to measure, PET being the easiest of all. Profile-based surrogate measures on the other hand consider vehicle trajectory and hence provide much more information on the driver behavior over a longer period of time. However, in-field data collection of such measures such as speed profile, acceleration-deceleration profiles, TET and TIT among others have not been considered owing to the difficulty in the measurement and the data intensity of these measures. These measures were either obtained either through simulation or using GPS-equipped vehicle traces.

The third and probably the most important is their relationship with crashes. Though many previous studies conclude that there is promise in the use of surrogates, they could not show definitive results (Gettman et al, 2008). Many studies often resulted in low correlation between surrogate measure and crashes and this is probably because of considering non-serious conflicts too (Williams, 1981). Section 2.2 throws more light on this aspect.

2.2 IMPORTANCE OF THRESHOLD

Though the literature and past research identified several measures that can potentially act as surrogates for safety, the exact relationship between them has not yet been consistently established. Past research (e.g. Perkins and Harris, 1967; Hauer, 1982; Sayed and Zein, 1999) has often relied on the understanding that the frequency of crashes and frequency of non-crash events have the following relation or its variants (Hauer and Garder, 1986; Saunier and Sayed, 2008; Tarko et al., 2009):

$$C = k.S \tag{2.3}$$

Where,

C = number of crashes

S = number of non-crash (or surrogate) events

k = constant

Hauer (1982) laid emphasis on the idea that the constant K in the above equation should not change across locations for the relation between conflicts and crashes to hold true. However, Hauer and Garder (1986) allowed limited variability in the value of K and extended the above equation to conflicts of multiple severities as

$$C = \sum_{i} Ki * Si \tag{2.4}$$

Where i denotes severity level

One of the reasons for inconsistent findings for the above relationship might be the definition of non-crash events. Considering all non-crash events without a proper threshold consideration makes the surrogate event synonymous to exposure, leaving the non-crash event (surrogate) adding little additional value. A study conducted by Zimmermann et al. (1977) found that the different degrees of severity of conflicts (slight, medium, and serious) are generally recorded in the ratio of 80:19:1, while this relationship depends on site and type of conflict. The surrogate should be such that it encompasses the risk of such an event leading to a crash, thereby having a value in addition to exposure (traffic volumes). Williams (1981) questioned the validity of TCT when he found no relationship between conflicts and crashes and he attributed it to the lack of standard operational conflict definition and combined counts of conflicts of various severities. Even assuming that the above relationship between near-misses or critical incidents and crashes holds true, the practical problem arises in the identification of these near misses out of an abundance of events. This alludes to the importance of identifying a threshold value to separate critical events from other safer events.

Past research shows that threshold identification has been mostly arbitrary and there have been no standard methodology or practice developed to identify a threshold. An inaccurate threshold might show a very weak relationship between the surrogate and crashes. Usage of an inaccurate threshold for a metric can weakens its predictive power but might still hold some value in categorization. For example, a surrogate might be able to tell if an intersection belongs to a high, medium, or low category of crashes but might not be able to predict the number of crashes that would occur next year. Usage of an accurate threshold might improve the correlation between surrogate and crashes. The statement by Baker and Glauz (1977) that “Traffic Conflict Technique was used mostly as a diagnostic tool as opposed to a predictive tool” alludes to the role of threshold in distinguishing the effectiveness of a surrogate as a predictive tool from a diagnostic tool. Hauer and Garder (1986) also argued that TCT should be used mainly as a diagnostic and evaluative tool rather than a predictive tool. Moreover, it is also possible that threshold varies from one surrogate measure to another and the threshold for the same surrogate might be different based on the type of conflict being studied. Therefore identifying an accurate threshold value for the non-crash event (surrogate) is extremely important to ascertain its capability in evaluating crash propensity of a location.

One of the earliest studies which mentioned a threshold for conflicts is Hyden (1977). Hyden observed from video-tapings that some of the road-users performed a safety maneuver as late in the interaction as possible which brought down the minimum "time to accident" (T_0-value) to as low as 1.5 sec. He states this as a threshold level for a serious conflict. However, contrary to the conclusions by Hyden (1977), they concluded that the best correlations of accident and conflicts are found for slight conflicts. They further observed that while approach roads show a well-established relation between conflicts and crashes, no such relationship was found for inner area of an intersection especially between left-turn accidents and left-turn conflicts.

Allen et al. (1978) proposed a definition of conflict as satisfying the two conditions of gap time and PET being lesser than their respective thresholds even though the vehicles are not necessarily on a collision path. However, he did not mention any threshold values in his study. Cooper (1976) conducted a study which involved the film examination of a single intersection and recording conflicts periodically over a period of one year. Results from this study showed that serious conflicts correlated better with accidents than those of non-serious nature and that PET had the highest correlation coefficient (about 0.5) with crashes. A study of merging vehicles at freeway entrances however concluded that even small gaps (1.5 seconds) accepted by merging vehicles have no correlation with accidents at that location. Sonchitruska and Tarko (2004) fitted a series of negative-binomial regression models to PET counts at thresholds varying from eight seconds to one second at 0.5-second spacing and found that a threshold of 6.5 seconds gave the best fitted models.

Elvik et al. (2009) attempted to derive elementary units of exposure as specific events that directly represent opportunities for crashes, in other words serious conflicts. They considered simultaneous arrivals from potentially conflicting directions of travel as an elementary exposure events and then go on to use a time interval of 1s assuming that it is short enough to be considered as a real potential conflict. However, no analysis or previous research studies were used to support this assumption. The Surrogate Safety Assessment Model (Gettman et al., 2008) uses two threshold values for surrogate measures of safety to identify conflicts among all vehicular interactions. SSAM utilizes a default TTC value of 1.5 seconds and PET value of 3 seconds which the analyst may optionally override with his or her preferred alternate values. Huang et al. (2012) tested the estimates provided from SSAM with field measured traffic conflicts. They used VISSIM to generate trajectories which in turn is used by SSAM to detect conflicts. A measure called mean absolute percent error (MAPE) was used to measure the difference between observed and simulated conflicts. To determine the threshold to be used in SSAM software, a range of thresholds were tested for two surrogates: TTC and PET. The range of TTC was set from 0.7s to 3.0s for rear-end conflicts, and the study found that the best goodness-of-fit and the lowest MAPE value was observed for a TTC threshold of 1.6s. Similarly, the optimum PET value was found to be 2s for which the lowest MAPE value was achieved for crossing conflicts. However, this study has shown that the performance of conflict prediction models for crossing and lane change conflicts was only moderate. In a simulation study by Archer and Young (2009) the safety effects of five different signal treatments were measured by two main safety indicators, one of which is conflict time-proximity or PET for which they assumed a threshold of 4 seconds. Though this was a before-after treatment study and not a study that involved evaluating the efficiency of PET as a surrogate, it shows the lack of proper guidance on PET threshold selection.

In summary, it can be seen that though there have been various efforts at identifying and using surrogate measures to evaluate safety, only a few tried to validate them. Moreover, the exact relationship between surrogate measures and crashes has not yet been consistently established which undermines the effectiveness of using surrogates. The lack of standard operating definition for a conflict, and a lack of knowledge about the threshold values of a surrogate that are used for identifying serious conflicts are the primary reasons for this lacuna in the usage of surrogate measures for safety evaluation. In most of the research efforts, the use of threshold values has been arbitrary. Though literature shows that there is some consensus on the use of a low threshold value (in the order of 1.5sec) for TTC, such consensus has not been seen for other surrogates. The only surrogate for which a threshold is thought about is PET though threshold usage has been arbitrary. In fact, as early as 1977 it was suggested that surrogate measures should be used as a diagnostic tool rather than a predictive one. Therefore, future research needs to work on the issue of establishing thresholds for various surrogates and conflict types, and understanding its power as a predictive tool vis-à-vis a diagnostic one.

The discussion until this point concerned the various surrogate measures identified in the literature. A detailed discussion about some of the studies which proposed surrogate measures, the data collection methodologies adopted, strengths and limitations of the surrogates measures and corresponding data collection methodologies was presented. Various studies were presented that disputed the idea of predictive power of conflicts, which instead suggested that conflicts and surrogates could be used as diagnostic tools. This was tied to the lack of a standard definition for a conflict and surrogate, and insufficient importance given to the idea of "threshold". The importance of a threshold value for a surrogate measure, its role in identifying serious conflicts among others and thereby in establishing the effectiveness of a surrogate was discussed. The literature review showed how the selection of threshold value, especially for PET was done in an arbitrary manner. There is significant value in establishing such threshold values for various surrogates and conflict types before trying to predict crashes using them.

Once a surrogate is identified and the corresponding data is collected, the next step would be to evaluate the relationship between the surrogate measure and crashes to establish the former's effectiveness. Therefore, the next logical step would be to review the various methods and statistical modeling techniques applied in previous studies to model crashes using surrogate measures or other factors of safety.

2.3 STATISTICAL MODELING

This section reviews the various statistical modeling techniques used previously by researchers to evaluate safety. As seen in the previous sections, there have been various studies that evaluated the relationship between surrogate data and crash history. Generally, the model forms which researchers use to model crashes using factors of safety are also used to model crashes using surrogate data. Therefore it is pertinent to first review those statistical techniques that were used to model crashes with, or without, surrogates as parameters. These models are thus called crash-based models. Exploiting the idea that near-crashes can be considered as surrogates for crash data, there have been a few studies that tried to find crash propensity by using surrogate data without taking crash history in the modeling process. Such technique is called non-crash based safety modeling. Development of reliable statistical models for estimating crashes or evaluating safety is very important for transportation safety studies. A discussion on both crash based and non-crash based methodologies for safety evaluation is presented here.

2.3.1 Crash based safety modeling

Crash based safety evaluation is still a very popular approach for safety analyses and the literature presents many statistical methods to model crash frequency. Most of the earlier studies have been devoted to establishing the relationship between crashes and traffic volumes, practitioners could directly see that locations having higher traffic volumes had higher numbers of crashes. Later studies began to include other perceived factors of safety such as geometric characteristics of the roadway and other traffic characteristics.

2.3.1.1 Additive and Multiplicative Models

Hauer (2004) summarized that there are three forms of statistical models commonly used for road safety research:

Additive model: $Y = L^*(\beta_0 + \beta_1X_1 + \beta_2X_2 + \ldots\ldots + \beta_nX_n)$ (2.4)

Multiplicative model: $Y = L^*(\beta_0X_1^{\beta 1}X_2^{\beta 2}X_n^{\beta n})$ (2.5)

Multiplicative model (exp. base): $Y = L^*\exp(\beta_0 + \beta_1X_1 + \beta_2X_2 + \ldots\ldots + \beta_nX_n)$ (2.6)

Where

L = the length of segment

Y = the expected number of crashes

X's = the covariates of the factors such as traffic volume, sight distance among others that affect safety

β's = the regression coefficients.

In case of intersections, the term L may be dropped. In addition, Hauer suggested a generic model:

$Y = \alpha^*L^*$(Multiplicative Portion + Additive Portion) (2.7)

Where,

Multiplicative Portion = f_0(AADT)*f_1(X1)*f_2(X2)… and

Additive Portion = g_1(AADT, X1')+g_2(AADT, X2')+…..

$f_0(.)$, $f_1(.)$, $f_2(.)$, $g_1(.)$, $g_2(.)$ denote functions.

Each covariate in the model might have a different form of relationship with crashes. Hauer suggests that an additive model corresponds to situations in which presence of a trait called point hazards such as driveways or a narrow bridge adds a certain number of crashes per unit length. On the other hand, measures such as AADT or shoulder width add to the probability of a crash instead of linearly adding to the number of crashes. Effects of such factors are explained through multiplicative models. Accordingly, X1, X2 etc. are factors that are expected to have multiplicative influence while X1', X2' etc. are expected to have additive influence. α is a shape parameter that takes into account factors that are not intersection or location specific but generic in nature such as driver characteristics, weather conditions etc.

David and Norman (1975) was one of the earliest known studies using additive model where they developed a linear regression model for crashes per 3 years at 82 intersections in San Francisco Bay Area. They found that traffic volume is the most significant factor followed interestingly by number of U-turn restrictions, and number of right turn lanes. One of the first found studies to use conflicting traffic volumes instead of direct AADTs was conducted by Hakkert and Mahalel (1978). They analyzed four-legged intersections in terms of 24 crossing or merging pairs of traffic flows where for each pair, they calculated the product of the two flows and then summed all 24 pairs of products to obtain a traffic flow index X for the intersection.

Various other studies also explored the use of additive models in the form of multiple linear regression (Resende and Benekohal, 1997 used volume-to-capacity ratio, medium width, and surface rating as parameters), and non-linear regression (Khan et al., 1999 used traffic volume, segment length, and vehicle miles traveled as parameters) models. Similarly, studies such as (Konduri and Sinha, 2002; Chatterjee, et al., 2003; Chueh, 1996; Caliendo et al., 2007) explored the use of additive models in developing crash predicting models.

Though additive models were explored in a few studies as mentioned above, multiplicative and mixed models are more popular as it is well understood that relationship between crashes and many factors of safety is non-linear. One of the earliest studies in this direction was conducted by McDonald (1953) in California where he developed multiplicative model for crashes. He studied 150 three-legged and four-legged intersections on divided highways, stop-controlled on minor legs and found the following relationship.

$$N = 0.000783*(V_d)^{0.455}*(V_c)^{0.633} \qquad (2.8)$$

Where N is the number of crashes per year, V_d and V_c are entering major road and minor road AADTs respectively. Another study in California by Webb (1955) developed three different models corresponding to two-phase signalized intersections in urban, semi-urban, and rural areas.

$$N_U = 0.000189(ADT1)^{0.55}(ADT2)^{0.55} \qquad (2.9)$$

$$N_S = 0.00389(ADT1)^{0.45}(ADT2)^{0.38} \qquad (2.10)$$

$$N_R = 0.00703(ADT1)^{0.51}(ADT2)^{0.29} \tag{2.11}$$

The three categories of areas of urban, semi-urban, and rural were differentiated by speed limits. Both the above studies considered only traffic volume in their models. Later studies started to evaluate the effect of other intersection characteristics on safety of intersections or road segments.

Turner and Nicholson (1998) tested various model forms using the relationship between crashes and traffic flows as an example, and concluded that nonlinear model forms performed better than linear models. Charles V Zegeer worked on various studies that explored the relationship between crashes and different factors of safety. Zegeer et al. (1981) explored the effect of lane and shoulder widths on accident reduction on rural two-lane roads. Zegeer et al. (1987) studied the effect of cross-section design on the safety of 2-lane roads. Similarly, studies such as Hauer et al. (1989) and Hauer et al. (2004) used multiplicative models to study safety at signalized intersections, and urban four-lane undivided road segments respectively. The multiplicative model was also used in studies such as McDonald (1966), Leong (1973), Alijanahi et al. (1999) among others.

The third model proposed by Hauer (2004) which is a multiplicative model with exponential base forms a subset of Generalized Linear Models, explained in section 2.3.1.2 in detail.

2.3.1.2 Generalized Linear Modeling

Generalized linear modeling (GLM) approach is currently the most frequently used technique to model crash counts. The typical form of the general linear regression model is

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots + \beta_n X_n \tag{2.12}$$

Where

Y is the response variable that is the estimated mean number of crashes

X's are the covariates representing site characteristics such as traffic volume, sight distance etc

β's are the regression coefficients

The GLM approach suggests that the actual estimated crashes at a location can be assumed to follow a separate distribution, and the mean number of crashes Y of this distribution is assumed to be related to the model of covariates through a link function. A GLM typically assumes this model of covariates to be of a linear nature. In safety studies, crashes are commonly assumed to follow a Poisson distribution (Fridstrom et al., 1995; Nelder et al., 1972) or a Negative Binomial distribution (Hauer et al., (1988)). Typically a link function would be exponential, square root, inverse or identity. More details about this methodology can be found in the "Analysis" chapter.

Pickering et al. (1986) considered crashes at three-legged intersections of two-lane roads, developed a Poisson model to predict the mean number of crashes per unit time, and found that products of conflicting volumes tend to be most significant. One of the properties of a Poisson distribution is that the mean is equal to the variance. But in many situations, this condition does not hold true due to the significant variation in the data. This variation is generally represented by a dispersion parameter which is assumed to have a Gamma distribution with parameters, say k and a. Hauer et al., (1988) was one of the first studies to suggest such a modification to the Poisson regression model where the response variable Y is distributed as a combination of Poisson and Gamma distributions resulting in a Negative Binomial distribution. Later, studies such as Dean et al. (1989) and Poch et al. (1996) used the negative binomial forms to model crash counts and found better correlation with crashes than that with Poisson regression. Comparisons have also been made by researchers such as Miaou et al. (1993), Bauer et al. (1996) and Vogt et al. (1998) and they suggest that negative binomial to be preferred if the data are sufficiently overdispersed.

Park et al. (2009) have used a variant of the traditional Poisson and NB regression called finite mixture models which are especially useful when the model is generated from heterogeneous data. Their results showed that the standard NB models fail to capture some important characteristics of the data especially when such data inherently has a mixture of multiple sub-populations having different characteristics (homogeneous data).

One of the limitations of GLM modeling technique is its inability to take into account correlations or data having time-series characteristics. Chin et al. (2003) used a Random Effect Negative Binomial (RENB) to identify geometric characteristics, traffic factors, and traffic control measures at signalized intersections in Singapore. However, they also acknowledge that the findings may be limited by the relatively small sample size. Some of the other parametric regression techniques used in the previous studies are multivariate Poisson-Lognormal regression (Ma et al. (2008), Karim et al. (2009)),

Parametric modeling and regression analysis techniques are popular in studies relating to surrogate measures. Parker and Zegeer (1989) found that the relationship between traffic conflicts and crashes is linear and statistically significant. They also compared the hourly conflicts with the sums and products of traffic volumes using simple linear regression. Djikstra et al. (2010) conducted a study in Netherlands where they modeled 300 km^2 of road network in PARAMICS. Conflicts were identified from the simulated model and GLM approach was used to develop models to predict crash frequency. The output from the model was compared with actual crash data for six years corresponding to 569 intersections in the road network. The study concluded that there was a significant relationship between observed crashes and simulated conflicts. Such approaches were also used in studies such as (Songchitruska and Tarko, 2006; Gettman et al., 2008, Boonsiripant et al., 2011).

2.3.1.3 Non-Parametric Modeling

Non-parametric regression provides an alternative approach to model crashes where there is no requirement to assume any distribution or parameters of the model. This is a very popular approach for studies that involve identifying critical factors, classification, or crash and injury severity. For example, Kuhnert et al. (2000) employed logistic regression, Classification And Regression Trees (CART) and Multivariate Adaptive Regression Splines (MARS) to analyze motor vehicle injury data. They demonstrated that CART and MARS have the capability to identify groups of independent variables, identify threshold values, and then classify the dependent variable. They further suggested that CART and MARS can be used as exploratory tools for a more detailed regression analysis. .Karlaftis and Golias (2002) applied hierarchical tree-based regression (HTBR) to analyze the effects of road geometry and traffic characteristics on accident rates for rural two-lane and multilane roads. Their analysis showed that the factors that are critical for determining incident rates at rural 2-lane roads and multilane highways are different. Similar studies were conducted using tree-based models to analyze accident rates and crash severities (Sohn and Shin, 2001; Sohn and Lee, 2003; Abdel-Aty et al., 2005; Park and Saccomanno, 2005, Pande et al., 2010, Abdel-Aty et al., 2010). Artificial neural network (ANN) is another non-parametric model frequently applied in studies relating to traffic safety (Mussone et al., 1999; Abdelwahab and Abdel-Aty, 2001; Abdel-Aty and Abdelwahab, 2004, Haleem and Abdel-Aty, 2012).

Recently there is growing interest in the use of non-parametric methods in studies relating to pre-crash maneuvers, driver behavior and "surrogate" measures. Some researchers (Hauer and Garder, 1986; Baker and Glauz, 1977) are of the opinion that surrogate measures should be used as a diagnostic tool rather than as a predictive one. They argued that a good correlation between surrogates and crashes is not necessary and it would suffice if the surrogate measures can reflect on the difference in safety rather than absolute values. As explained in section 2.2, the statement by Baker and Glauz (1977) that "Traffic Conflict Technique was used mostly as a diagnostic tool as opposed to a predictive tool" alludes to the fact that conflicts recognized by the TCT method did not show a very strong correlation with crashes so as to have predictive power. There is value in studying if this argument can be extended to other surrogates too. The relationship between threshold value of a surrogate and its capability to act as a predictive tool vs. diagnostic tool needs to be investigated. Such thresholds, which can later be used to diagnose safety of a location, can be established by using non-parametric methods such as tree-based regression as a first step (Kuhnert et al., 2000). In this way, a location can be diagnosed as belonging to a safety category (for example, high, medium or low) without actually predicting crash frequency.

Kim (2006) in a study concluded that total crash numbers do not reveal traffic conditions or geometric variables related to each crash-type and then extended that study (Kim et al., 2007) where they used data from rural intersections in Georgia to model crash types using hierarchical multilevel modeling approach. They showed that crash data in Georgia in fact shows hierarchical structure: drivers' characteristics are nested within crashes, crash characteristics are nested within site characteristics, and site characteristics are nested within regional characteristics. This is a significant observation as crash prediction models are generally based on site characteristics and they are in turn assumed to influence driver's characteristics. But the results of this paper support the case for use of surrogates in crash prediction models as most of the surrogates are a measure of driver behavior that results either from site or driver characteristics.

Harb et al. (2009) used decision trees along with random forests to analyze P_CRASH3 (a binary variable that identifies whether a driver performed evasive actions prior to the accident occurrence or not) for rear-end, head-on and angle collisions. This technique ranked the various factors of intersection safety including driver and vehicular characteristics while eliminating potential correlation effects. The authors however acknowledge that the value of the main variable P_CRASH3 is judged by the police officer at the site of crash based on evidence and hence is a limitation. They also admit that due to limited sample size, the study could not differentiate between the different evasive actions, driver distraction causes and visibility obstructions.

Other recent studies that used non-parametric analysis methods that dealt with surrogate safety measures are Gettman et al. (2008), and Boonsiripant et al. (2011), Wu and Jovanis (2012), Hossain and Muromachi (2013) among others.

2.3.1.4 Empirical Bayes (EB) method

Another approach to crash prediction which gained popularity in more recent years is the Empirical Bayes(ian) (EB) method as it takes into account both the historical crash information of the study location and the estimated crashes for a population of locations similar to the study site. In fact this method supplements the various statistical modeling techniques discussed in section 2.3.1.3. In 1980, Hauer proposed a very simple method for estimating crashes, based on the knowledge of the distribution of crashes in a population of study units (Hauer, 1980). This idea led to the development of the EB method. The EB methods helps in controlling for the effect of regression-to-mean bias while considering the actual number of crashes at any location and thereby helps to increase confidence in using such a number. This method uses a combination of two measures to estimate the safety of an entity: the crash history of that entity and the crash frequency expected at similar entities. The expected crash frequency at similar entities is determined by the Safety Performance Function (SPF). The final estimated crash frequency at the entity is a weighted average of the two aforementioned measures.

$$N_{EB} = w*N_{SPF} + (1-w)*N_{OBS} \tag{2.13}$$

Where

N_{EB} = Expected accidents for an entity by EB method

N_{SPF} = Expected crash frequency at similar entities

N_{OBS} = Crash frequency observed at the entity

w = weight given to accidents expected on similar entities

“Weight” will depend on the reliability of both the measures from which the final expected crash frequency is estimated. SPFs are calibrated from data by statistical techniques such as GLMs explained before. SPF gives an estimate of the average number of accidents as a function of regression parameters (such as AADT, intersection characteristics etc.). This function is developed using locations having similar characteristics as the entity under study. Negative Binomial regression is currently a common approach to build SPFs and one of the parameters of this distribution is the “overdispersion”. The value of this parameter is required to calculate “weight” used in EB method.

$$\text{Weight} = 1/(1+(N_{SPF})/\Phi) \tag{2.14}$$

Where Φ = overdispersion parameter

Belanger (1994) estimated the safety of four-legged unsignalized intersections using the Empirical Bayesian method. SPFs for the population of similar intersections were developed using major and minor flows and the best model was obtained from the product of major and minor flows. The results were used to identify hazardous locations and evaluate effectiveness of treatments. Persaud and Nguyen (1998) calibrated two levels of models to act as SPFs based on the EB approach. The first level of models estimated crash frequency as a function of the total intersection entering volume. The second level of modeling attempted to disaggregate the modeling based on the type of collision by considering flows pertinent to type of conflict. The estimation of safety using the Empirical Bayesian approach has also been investigated by several other researchers (Abbess et al. (1981), Hauer et al. (1986), Hauer (1997) and Garber et al. (2001)).

2.3.2 Non-crash based safety modeling

This section discusses the techniques used in previous studies to predict crash propensity without actually using crash history to build the model. The literature relating to the identification and data collection methodologies of the various surrogate measures has already been discussed in the section 2.1. While this modeling is non-crash based, crash data is important for validating the non-crash based models as crash data is the direct measure of safety. Hence it can be said that non-crash based models are not completely independent of crash data. Statistical modeling techniques for non-cash based safety evaluation are not as extensively developed as for their crash based counterparts. One of the reasons for this is that the relationship between surrogates and safety is not as clearly understood as that between crashes and safety. The reasons for this lacuna were explained in section 2.2.

Songchitruksa et al. (2006) used an approach called “extreme value technique” to examine the validity of PET as a safety indicator. Extreme value theory (EVT) is a statistical method used to evaluate the risk of rare events such as floods, earthquakes etc. A PET value of 0 implies a crash and a crash can be considered to be a rare event given the number of opportunities that exist in terms of traffic volumes or exposure. Hence an EVT method was used to estimate the probability of the occurrence of a PET value of 0.

This approach was used to estimate probability of right angle crashes (between through vehicles of major and minor approaches). Although EVT has been used in many applications in multidisciplinary areas its first application to transportation problems was found in the work by Hyde and Wright (1986) (also corroborated by Tarko (2012)) where they used extreme value methods to estimate road traffic capacity. However, application of this theory to transportation safety is found in the study conducted by Campbell et al. (1996)) who applied Weibull distribution to traffic conflicts to evaluate the benefits of active safety technologies. Chin et.al. (1991, 1992) analyzed safety of expressway merging interactions by collecting time to conflict information from video recordings. The data was fitted using a Weibull distribution, the tail of which gives the probability of accidents. Recently, Oh et al. (2010) derived the crash probability by using an exponential decay function using time-to-collision (TTC) between two vehicles. Tarko (2012) used generalized Pareto distribution to find probability of departure crashes.

2.3.3 Section Summary

The statistical methods generally used to model crash frequency can be broadly classified into three categories: (a) Parametric regression (b) Non-parametric regression and (c) Bayesian methods. In safety literature especially dealing with surrogates, the statistical methods can also be divided as crash based and non-crash based methods. Crash based methods consider incident data in the modeling process but non-crash based methods look to estimate crash propensity directly using some distribution of surrogate data and use crash data only to validate the prediction. Parametric regression using generalized linear models (GLM) is the most widely used method in safety modeling, though many new statistical techniques are being explored for a variety of special situations. Non-parametric methods also show lot of potential, especially for qualitative predictions (such as categorizing safety at a location, modeling severity of crashes).

Though non-crash based methods have shown promise, their applicability is limited because in the absence of crash data, the non-crash measure that is used should represent a crash at some value. For example, since a PET value of 0 implies a crash, the probability of occurrence of a PET value of 0 will imply a crash. Therefore PET data distribution can somehow be used to predict crashes, without actually using crash history. But this might not hold true for all surrogates. TTC on the other hand is an “expected” time that does not take into account evasive maneuver, and extrapolating TTC distribution to find probability of crash would lead to over-estimating crash frequencies. Moreover, as of now, only certain statistical distributions such as extreme value distributions are found to be applicable for such safety evaluations. Therefore, crash based modeling seems to be the better approach of the two.

2.4 IDENTIFICATION OF STUDY SITES

The last aspect of the analysis of the existing literature deals with methods for selecting candidate intersections for the study. Hot spot identification is an essential when it is necessary to identify locations with high levels of incidents, such as roadway safety improvement projects. Ineffective identification of hot-spots can lead to wastage of resources or preserving unsafe conditions. The research discussed in this study will use hot spot identification techniques in the study of PET. The candidate intersections selected for evaluating the effectiveness of PET should not be biased and should have a mix of intersections that have high, medium and low level of safety and traffic volumes. So, the task of identifying these intersections is essentially a task of ranking a set of intersections with respect to safety by selecting the appropriate method of hot-spot identification.

The simplest way of finding the least safe locations is to rank them according to crash count or frequency (Deacon et al., (1975)). As already explained in section 1.1 as a limitation with using crash data, this method will result in regression-to-mean bias due to the randomness inherent in the accident counts. This method also tends to be biased towards locations having higher traffic volumes as they tend to have more crashes and this higher number of crashes may be just due to greater exposure of vehicles rather than the general safety level of the location. The second method which takes into account both traffic volume and crash frequency into account is ranking by accident rate but this also suffers from regression-to-mean bias. To correct for this bias associated with the typical hot spot identification methods, Empirical Bayesian techniques (EB) are used. The EB method was originally developed in order to control the regression-to-mean bias in the before-and-after studies evaluating the effects of roadway safety treatments (Hauer et al., (1983), Persaud et al., (1984), Hauer, (1997), Persaud et al., (2007)). This method has also being used for hot-spot identification (Persaud et al., (1999), Miaou et al., (2003), Cheng et al., (2005), Elvik, (2007)). The validity of the EB approach has also been establishes by multiple studies (Persaud et al., (1984), Hauer et al., (1983), Elvik, (2008)). Variants of the above discussed methods such as equivalent property damage only (EPDO) crash frequency (Montella, 2009), proportion method (Lyon et al., 2007), Empirical Bayes estimate of severe crash frequency (EBs), and potential for improvement (PI) (Persaud et al., 1999; El-Basyouny and Sayed, 2006) have also been evaluated in various other studies.

A Bayesian updating reliability method (Haleem et al. (2010) was also used to update the parameter estimates of covariates in a NB model to improve the crash prediction accuracy of such models developed for 3-legged and 4-legged unsignalized intersections. This study concluded that though both NB model and Bayesian models predicted crashes alike, the Bayesian updating framework using the log-gamma likelihood function for updating parameter estimates of the NB models resulted in the least standard error value which the study considered as a surrogate for uncertainty. Several works have also compared the various hot-spot identification methodologies (Persaud et al., (1999), Montella, (2009), Cheng et al., (2005, 2008), Elvik, (2007, 2008)). The results from these studies demonstrated that EB method performs better than the other hot-spot identification methods.

This section of literature review tells us that care should be taken while selecting intersections based on crash numbers because crash numbers can be subject to regression to mean bias and there can be an additional bias with locations having higher traffic volumes. One of the ways to deal with the problem of regression to mean bias is to consider higher number of years of crash data (Nicholson, 1985) while another solution is to use Empirical Bayes method. EB method though is complicated and has its own limitations (Huang et al., 2009). Coming to the problem of bias due to traffic volumes, one can always divide the whole range of traffic volumes into groups having smaller ranges and then rank within the groups. But overall, selection of study locations plays an important role in the final result obtained from the study and hence should be done carefully.

2.5 CHAPTER SUMMARY

The use of surrogate safety measures is expected to allow for more rapid and earlier safety analysis relative to evaluations using actual crash data. Most of the surrogate measures proposed in the previous studies are traffic operational characteristics as all these are related to traffic and its events. These measures can be first categorized as macroscopic or microscopic. Many of the macroscopic measures include fairly standard measures of effectiveness. However, as stated by Gettman and Head (2003), these measures are used as “rule-of-thumb” for evaluating safety rather than for crash predictions. Microscopic measures on the other hand are more popular as surrogate measures as these directly measure individual traffic events. Microscopic measures can further be classified as point-based measures or profile-based measures.

Profile-based measures such as speed variation, acceleration noise, acceleration deceleration profile, TIT and TET have not been explored much due to the difficulty in measuring them. These measures require obtaining the trajectory of the vehicle and robust automated vehicle tracking systems have not been fully developed yet especially for dense traffic conditions. Therefore, the only way to obtain these measures is either by simulation or from GPS equipped vehicles. Point-based measures on the other hand are easier to measure. Various studies have evaluated some of these measures and literature shows that surrogate measures of safety at best provide mixed conclusions about their effectiveness and applicability.

Many studies often resulted in low correlation between surrogate measure and crashes. One of the primary reasons for this is that the definition of conflict and criteria for identifying a surrogate measure are still a matter of debate. Moreover, there is a lack of understanding and research about the method to recognize serious conflicts among a mixture of serious and non-serious conflicts. A major requirement for filling this gap in knowledge is to establish a threshold value for a surrogate measure and using this threshold to differentiate serious conflicts from the others. The literature shows that selection of a threshold for most of the surrogates has been arbitrary, except for TTC for which there is a consensus among researchers that the threshold lies at 1.5 sec. Establishing such threshold values for other surrogates and conflict types is of high importance before using surrogate measures for safety evaluations.

Even with arbitrary selection of threshold values, researchers have used various statistical techniques to model the relationship between surrogates and crashes. Most of these models are borrowed from studies that evaluated the relationship between “factors of safety” and crashes. GLM approach with Poisson and Negative Binomial family is the most popular form of modeling such relationship. However, recently there have been a few studies that tried to estimate the crash propensity directly from surrogate data distribution. These are called non-crash based methods. Researchers in the past (Baker and Glauz, 1977; Hauer and Garder, 1986) have argued that TCT should be used mainly as a diagnostic and evaluative tool rather than a predictive tool. There is value in studying if this argument can be extended to other surrogates too. The relationship between threshold value of a surrogate and its capability to act as a predictive tool vs. diagnostic tool needs to be investigated. The SSAM study (Gettman et al., 2008) also evaluated the diagnostic capability of surrogate measures by studying the rank correlation between conflicts and crashes. Non-parametric methods such as tree-based regression techniques are generally used for such qualitative safety evaluations and this method shows promise for determining the diagnostic capacity of surrogate measures.

Finally, selection of study locations is also an important step in safety studies. When performing such selection using crash histories of locations, care should be taken so that the selection is not subject to regression to mean bias. Moreover, selection of locations based on crash frequencies or crash rates have biases in terms of traffic volumes. Though Empirical Bayes method is one method that can limit the regression to mean bias, it is a complicated method. Therefore, depending on the requirement, appropriate method of selection should be adopted such that there is as minimal bias as possible in the final selection.

CHAPTER 3: PHASE 1

3.1 INTRODUCTION

The overall idea of this dissertation is to evaluate the effectiveness of certain surrogate measures. For, the first step is to collect statistically sufficient surrogate data by developing a surrogate data collection methodology and then use the collected data to perform analysis and investigate the effectiveness and applicability of those surrogates. Phase 1 of this research represents the initial case study that was aimed at developing such a system and identifying the various challenges and requirements involved in it. This phase of the study was conducted to understand how a surrogate study could be performed, and what the issues are with potential data collection methods and analysis procedures. The literature review has shown that there are data collection issues especially with respect to profile based surrogate measures. On-field data collection of such surrogate measures have not been attempted much owing to the difficulty in such methods (this phase of research was conducted in the year 2008). Though there is recent growing application of computer vision techniques for automated vehicle tracking, at the time of this phase of research, such systems have not be implemented and tested on field at a scale that is widely applicable. Nevertheless, there are still limitations with respect to system requirements, and traffic conditions for such video detection systems. For example, one of the recent studies that successfully implemented such systems (Aubin et al. (2013)) has limitations in the case of dense and turbulent traffic flows and hence their methodology targets high-speed, low to medium-flow scenarios only. Moreover, a longer continuous trajectory would require multiple synchronized cameras viewing longer stretches of road potentially requiring perspective angles leading to further occlusion by vehicles in denser traffic conditions. In view of these limitations, this phase of research developed a data collection system that is semi-automatic in nature, can be used to collect both profile-based and point-based surrogate measures, and limits the system requirements and traffic condition restrictions by using human observers to track vehicles and record their trajectories. This chapter shows how the experiences from the data collection process, and analysis of the collected data led to the next phase of the research.

To accomplish the objectives of this initial phase of the overall study, the developed data collection system was used to collect surrogate data and then use it to evaluate the effect of a treatment on a high speed rural multi-lane highway intersection. This study was sponsored by GDOT and the intersection of US23/SR 365 & CR387 (Demorest MT. Airy Hwy) was selected by the sponsor. This intersection is located in Habersham County, Georgia. The study area is marked in the Figure 3.1 below.

Figure 3.1: Map showing locations of the study intersection with reference to the state of Georgia in the inset (www.mapquest.com)

This location was selected because the prior crash history at that location has shown safety concern with the interactions between left turn vehicles during permissive phase and opposing through vehicles. Though this is a rural intersection, a school in the vicinity of intersection creates significant number of left-turn vehicles, especially during the school timings. There is significant through vehicle traffic too, especially during the peak hours. The intersection being a high-speed intersection made the situation more precarious.

Before going into the crash history details, the operational and geometric summaries of the study intersection such as traffic control parameters, volume, sight distance etc. were prepared. Then the crash data for the intersection was reviewed and collision diagram was drawn. Crash data for a period of 3 years (2002, 2003 and 2004) was summarized by reviewing the incident reports. Figure 3.2 shows the collision diagram for the study intersection.

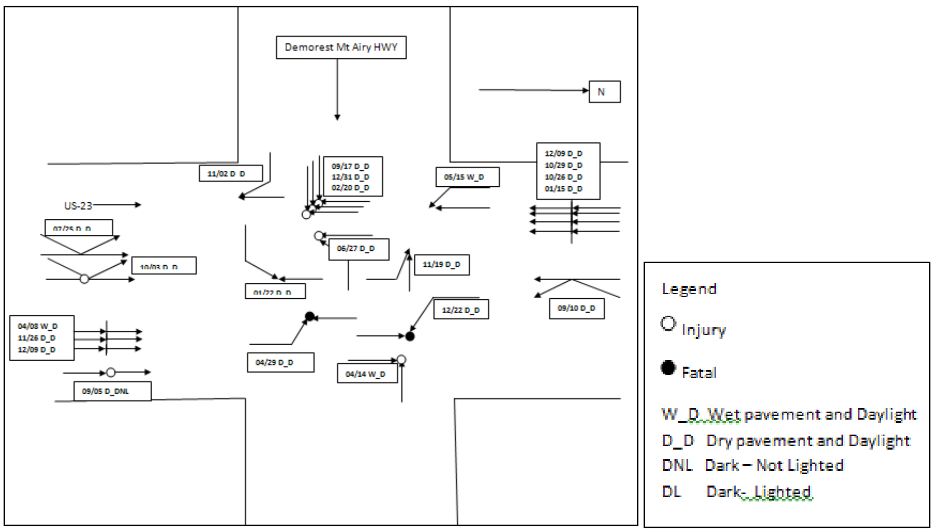


Figure 3.2: Collision diagram for the study intersection (crash data 2002-4).

The collision diagram not only acts as a summary for the intersection but also help in determining the critical conflict which needs to be addressed. The data collection methodology and surrogate measures which need to be captured are determined by the critical conflict. For example, from Figure 3.2, it can be seen that there have been two fatal accidents between a left-turning vehicle and an opposing through vehicle at the study intersection. Though the number of rear end accidents is more, the crash between a left-turn vehicle and an opposing through vehicle is more severe.

Based on the collision diagrams and discussion with GDOT, the critical conflict was identified as the opposing left-turn conflict. Moreover, it was also decided by GDOT that the treatment (Figure 3.3) would be a non-standard striping (i.e. a combination of 3 chevron stripes approximately 1000 ft upstream of the intersection and 3 stripes preceding the stop bar to serve as graduated stop bars).

Figure 3.3: (a) Chevron markings 1000 feet upstream of the intersection (b) Non-standard striping (in orange color) preceding the stop bar acting as graduated stop bars

The next prime task was to identify the surrogate measures that would reflect the interactions between a left-turn vehicle and opposing through vehicle. First of all, the conflict under consideration is of crossing nature. Most of the previous works (Allen et al., 1978; Cooper and Ferguson, 1976; Gettman et al., 2003) considered measures such as gap time and PET to be appropriate because they involve a common area of conflict between the crossing maneuvers. Cooper and Ferguson (1976) concluded that PET had better correlation with crashes than other measures considered. PET is a measure of an actual proximity to crash while gap time is an expected proximity. Therefore it was decided to consider PET as one of the surrogates for this study. Measures like time-to-conflict and its variations have always been used in case of rear end conflicts and merging conflicts where both vehicles have same direction of movement (examples include Hayward, 1972; Minderhoud and Bovy, 2001). It was also explained in chapter 1 with reference to the paper (Tarko et al., 2009) that a good surrogate measure would also capture the effect of a treatment. The treatment is expected to modify the behavior of through vehicles by making them slow down appropriately when they see a graduated stop bar (which is a few feet upstream of the actual stop bar) and a vehicle turning left at the intersection. Since this requires capturing the behavior of individual through vehicle, the surrogate measure considered should be such that it captures individual vehicle behavior. Braking, swerving, speed, acceleration noise, and deceleration rate are a few such measures. Previous studies have already shown the limitations and subjectivity involved with the observation of braking and swerving. These can be objectively measure with the measures of high accelerations or decelerations. So, acceleration deceleration profile could be one potential surrogate. Regarding speed, one of the points of concern is where to collect this information. It was decided to collect speed value of the through vehicle as it enters the intersection as this gives a final “result” from the interaction and the effect of treatment, if any.

Following the above thought process, it was decided that based on the critical conflict identified, and the reviewed literature on surrogate measures, the measures considered for this study were acceleration deceleration profile of the through vehicle, intersection entering speed of the through vehicle, and the Post Encroachment Time (PET).

3.2 DATA COLLECTION

The primary method for collection of surrogate data was decided to be a combination of video recording and post processing as videos will allow for a permanent record of the intersection operations, and the reduction of video to obtain surrogate data can be performed in laboratory environment.

3.2.1 Features of the equipment

In order to collect video data, a portable data collection station was developed (Figure 3.4). Each portable station consisted of a trailer equipped with solar panels that charge a set of six deep cycle marine batteries which supply a constant 12Volt DC power. The data collection unit was equipped with a pan-tilt-zoom (PTZ) network camera that could either be mounted on the mast of the trailer or mounted on a separate pole adjacent to the trailer. The use of a network camera, instead of an analog camera, allows for a direct connection of the camera to a low power notebook computer. The video stream was recorded on the notebook and periodically exported to an external hard drive. The setup also featured a wireless cellular network connection to the notebook that allowed a user to remotely control the camera (change the view, turn off/on etc.) as well as control the recording process. The video data collection station was designed with the objective of providing high flexibility in the communication, recording, and camera control processes on a low power budget, allowing solar panels to act as the primary external charging source for an extended period.

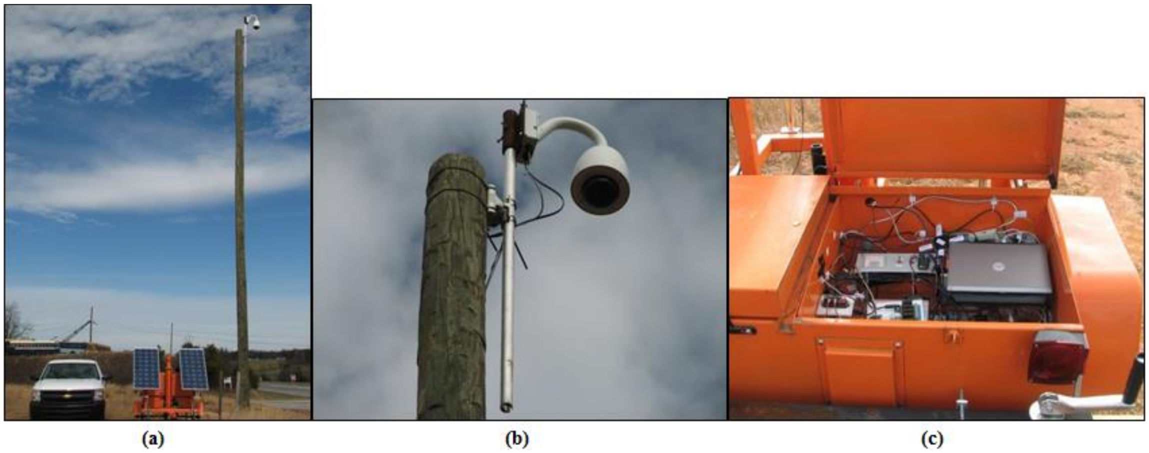


Figure 3.4: (a) equipment trailer at base of pole, (b) pole and the camera mounted on it, (c) trailer on-board equipment (Photo credit: Guin, A. (2008))

3.2.2 Pre-deployment experimentation

Prior to field deployment of the equipment at the actual intersection, it was tested thoroughly with respect to various important aspects that would affect the quality of the data collected. First of all, the height at which the camera is to be fixed needed to be determined as the viewing area and angle depend largely on this factor. So, a camera height test was conducted at the Georgia Tech Structures Laboratory to determine the optimal height of the camera. The portable data collection station had an in-build mast that can be extended up to a height of 15’ and the camera could be fixed to the end of the mast. However, in order to test the camera for greater heights, a bucket truck was requested from GDOT that could extend as high as 70 ft. With the help of this truck, the camera was tested and video clips obtained at heights ranging from 20 ft to 70 ft above the ground, in 10 ft increments.

Most of the automated vehicle detection or tracking systems such as NGSIM have cameras looking at the road with very high angles making the section of roadway under the view of the camera smaller. The profile-based surrogate measure being considered in this study requires recording the vehicle movement over a long distance of its approach to the intersection. So, a camera would be required to look at a longer stretch of road making the angle of view quite lesser creating a perspective effect. Given the study intersection, it was likely necessary to collect data from within a few feet of the roadside (although outside of any required clear zone), resulting in a relative shallow angle between the camera and portions of the roadway. That is, the camera had to be of sufficient height to allow for distinguishing between vehicles hundreds of feet from the camera, at a generally straight-on angle.

The second aspect of importance was the ability to observe detection lines or some on-ground features in the recorded video that would assist in getting the trajectory of vehicle. Initially, orange colored tape strips were placed on the road at increasing distance from the camera but were too small and could not be noticed in the video. The video was recorded again by placing cones at either sides of the road and this time the cones could be seen in the video. This experiment showed that cones can be placed in the field too that would assist in drawing detection lines.

The third aspect of concern was wind at the height where the camera is located. Even small movements of the camera due to wind may magnify the distortions in the recorded video because the camera is located far from the area under view. So, it needs to be ensured that the camera is as stable as possible. The camera height test showed that a height between 40 ft to 50 ft provided a good balance between visibility of vehicles and detection points, and availability of sufficient coverage area.

3.2.3 Field Deployment

The study intersection is on a high-speed rural multilane highway with a speed limit of 65 mph (104.5 km/hr). The length of the intersection approach that needed to be covered by the video was approximately 900 feet (274.32 m). This is the distance of the approach to the intersection along which the trajectories of the through vehicles were supposed to be recorded to get the acceleration deceleration profile. This distance was selected to exceed the distance required for a vehicle to stop based on stopping sight distance criteria for the given speed limit. This zone is also selected such that the vehicles enter the detection zone after they cross the chevron pavement markings 1000 ft upstream of the intersection (a part of the treatment that was supposed to be applied at the intersection).

Analysis of the video clips obtained from the camera height test and also some videos recorded at the field on a test basis revealed that a minimum height of approximately 40 ft to 50 ft is required to ensure clarity of the video. Lower height resulted in potential occlusion and difficulty identifying the front and rear of each vehicle. However, even given this height it was clear that a single camera is not sufficient to capture the video over a 900 ft zone, the estimated range required for the data collection. The maximum useable range was 600ft, and less where the roadway slopes down away from the camera location. It was therefore concluded that a two camera solution was necessary, one for data collection through the intersection proper and immediately upstream of the intersection and a second farther upstream to monitor approaching traffic. The possibility of running two videos from two cameras in synchronized mode in the GT video analysis software was evaluated. It was concluded that such architecture was theoretically feasible.

In consultation with GDOT personnel it was determined to utilize permanent pole placements at the Demorest-Mt. Airy Intersection for the video camera placement, as seen in Figure 3.5. A site visit with GDOT to determine optimal pole locations at the intersection of US23/SR 365 & CR387 (Demorest-Mt. Airy) was conducted. It was determined that four (4) 60 ft poles will be installed near the intersection. Though more cameras would mean that each camera would have smaller area to focus on and hence greater clarity in recognizing the vehicles and detection lines, it would also mean requiring installation of more poles, more equipment requirements, and more importantly more difficulties synchronizing and managing that many cameras looking at the same vehicle proceeding through different segments of the approach road. Therefore a decision was taken to have only two cameras looking at one approach for this study.

Final installation included two poles at 300 ft on either side of the intersection to capture a view of the left turn bay and the oncoming traffic and two poles approximately 1100 ft from the intersection to capture the upstream traffic approaching the intersection (Figure 3.5). An image of the intersection for the camera on the pole located approximately 300 ft south of the intersection is seen in Figure 3.3 (b). For this research effort the cameras were mounted on these wood poles at a height of approximately 45'. Strong wooden poles also provide stability to the camera against wind at that height. The camera views are overlapped to enable accurate vehicles identification and video synchronization during the post-processing of the multiple camera views.

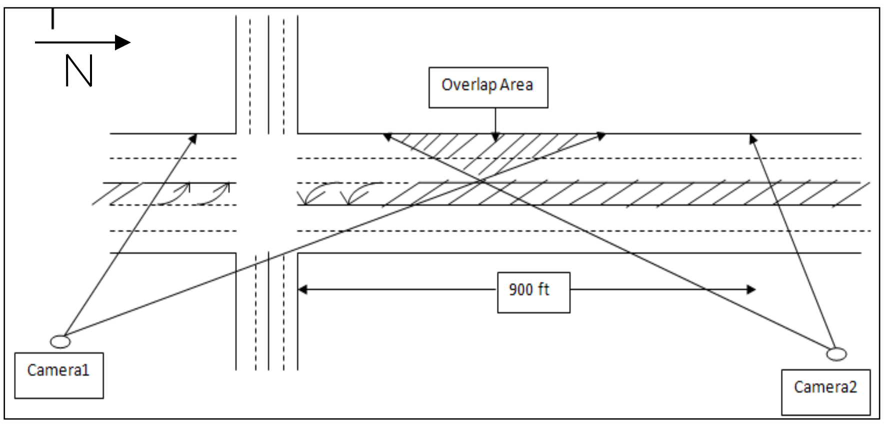


Figure 3.5: Field placement of the two cameras for the southbound approach

3.2.4 Issues with data collection setup

The previous section explained the components and setup of the proposed data collection system. Though there are advantages of such a methodology in terms of collecting on field profile-based surrogate data, there are certain impediments to the data collection process and it is important to recognize them.

- The power source for this portable system is a group of batteries that are charged through solar panels. Inclement weather conditions were an inhibitor during the data collection process. During the data collection in this phase of research, there were several long periods of cloudiness which hampered the charging of the batteries from the solar cells and created power shortage issues for the equipment in the data collection trailers.

- Placement of the trailers which had these batteries and other equipment was another point of concern. Positioning the trailers so as to avoid shadows from nearby trees while keeping the trailer out of the clear zone was important. This was solved by positioning the trailer in a sunny spot away from camera pole. However, that caused increased power losses in the extended power cables leading to more stress on the power adapters and the eventual damage to the adapter which could only be resolved with new adapters, better heat dissipation from the adapters and an upgrade of the power cables.
- Wind was another issue sometimes which caused the movement of camera away from the viewing area. Even though the recording continues from both the cameras, if one of the camera is out of vie, then the recordings from both cameras would be unusable as videos from both cameras are required to get the complete trajectory or profile of the vehicle.
- The camera was mounted on a wooden pole at a height of approximately 45’. Any problem with the camera itself required involving GDOT, asking its personnel to come to the field with a bucket truck to bring down the camera to replace it or fix it.
- There were other intermittent issues such as filling out of external memory drives, drop in the network signal causing breaks in the video records and difficulties with remote access of the portable system and camera.

3.2.5 Data Reduction Software Design

Custom software was developed in Java and using the Java Media Framework (JMF) technology to allow for a frame-by-frame review of the video. For any frame selected by an analyst, the software may be used to extract the frame number and timestamp. The custom software has two primary components – SaveGrid and ExtractData. SaveGrid allows the analyst to construct a video overlay containing detection lines separated by a set distance, 40ft in this effort, based on known locations in the field of view. The known points were determined as part of the initial field survey during the in-field equipment setup. The spacing of 40ft was chosen because the lane dividing markings on the roadway are spaced at 40ft (length of each line being 10ft and gap being 30 ft). A field test was done to check the accuracy of their spacing. Though they were fairly accurately spaced with a maximum error of +/- 1 ft, it was decided to place other identification markings at 40 ft spacing. It was found in the camera height test experiment that cones placed along the roadway could be detected in the recorded video and hence the same idea was implemented in the actual site. Cones were placed along the approach road at an interval of 40ft. Detection lines were then drawn connecting pairs of cones on opposite sides of roadway using SaveGrid component of the software. Given the accuracy of the spacing between the lane dividing markings, and given the resolution of the video, the lines drawn between cones coinciding with the lane divided markings. This overlay is saved and is later re-loaded in the ExtractData module for extraction of data from the video. The red detection lines in Figure 3.6 represent a typical video overlay.

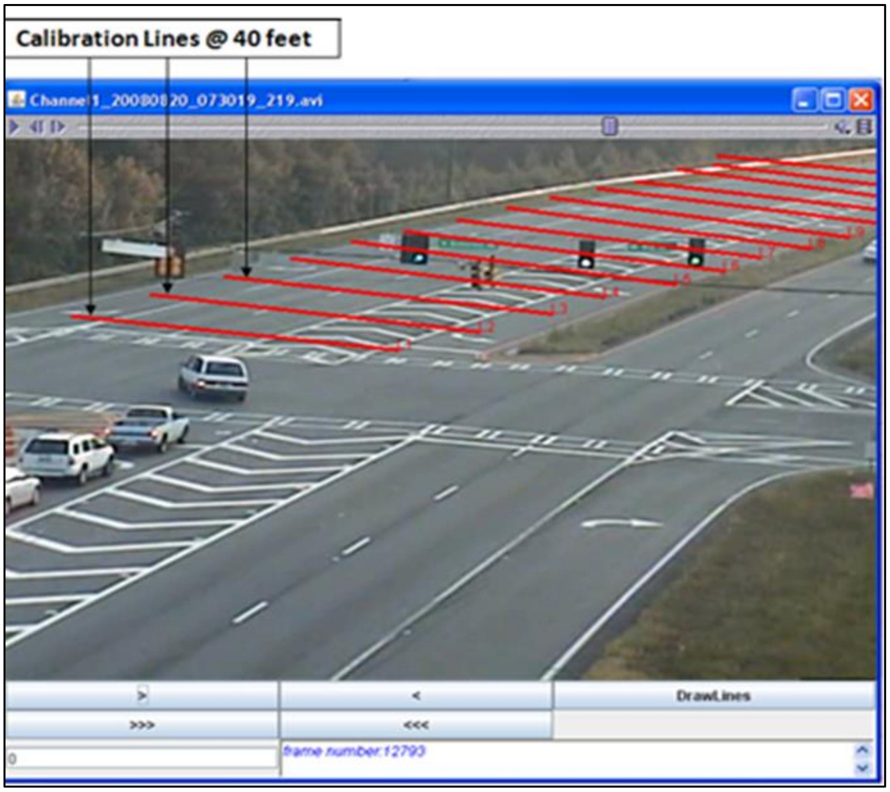
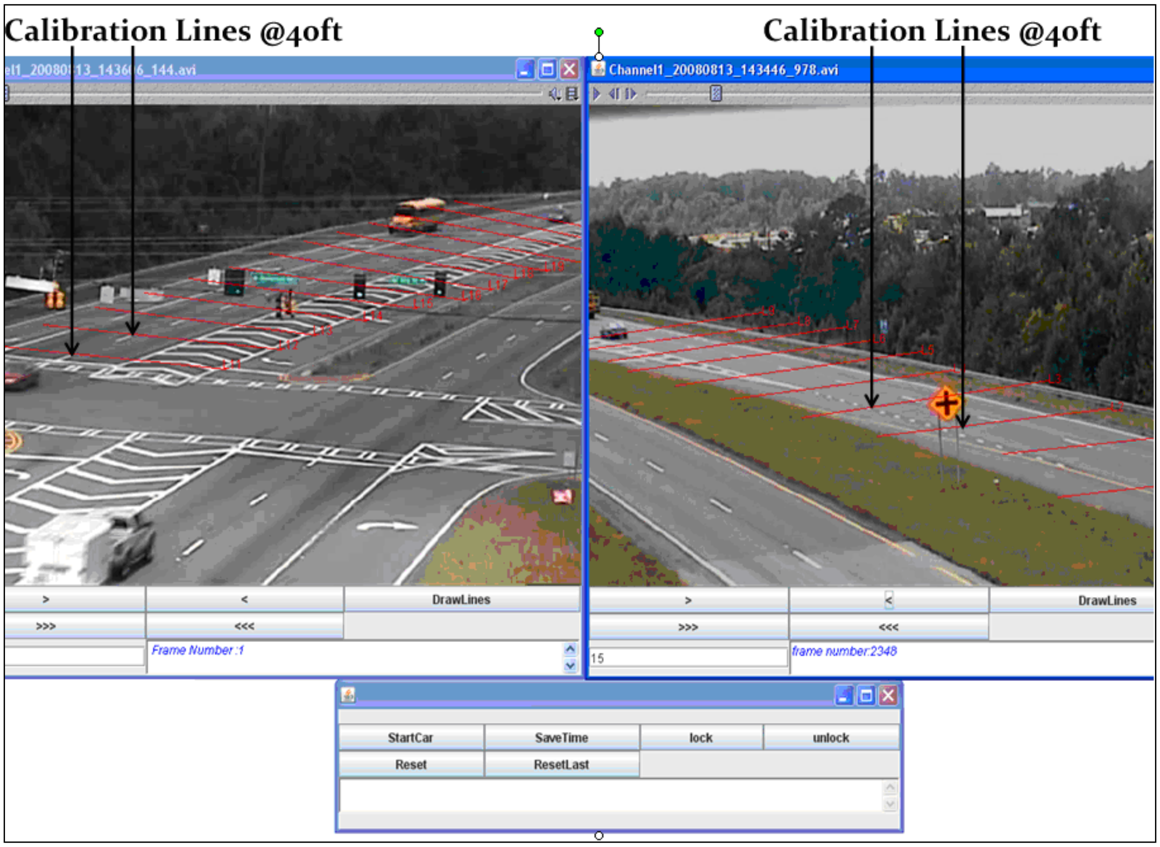


Figure 3.6: Example Screenshot of Video Reduction Software

ExtractData is used to collect the data from the video. A frame of the ExtractData component is shown in Figure 3.7. The software allows the data analyst to step through the video frame-by-frame (forward and reverse) as well as by a customizable multi-frame step for faster navigation. At the start of the data reduction the analyst imports the overlay detection lines created from the SaveGrid component. The analyst then extracts the time and position data of each vehicle as it crosses each detection line drawn in the video overlay. Using frame-by-frame (or multi-frame) to step through the video the analyst selects the frame in which the front tires of the subject vehicle are positioned on the detection line of interest. The analyst then selects the “Savetime” button. The distance of the vehicle from the stop bar (calculated using the overlay detection line number), the corresponding frame number, and the timestamp are recorded. The analyst also has several reset options to aid in handling analyst errors (such as saving the information for the incorrect frame or skipping a detection line while processing a vehicle). To minimize errors the analysts tracks one vehicle through the entire intersection approach prior to collecting data for the next vehicle. All data for a video is stored in a comma-separated-value ASCII file.

As stated previously, two cameras are used to capture each intersection approach. Figure 3.7 provides an example of the video from two cameras for an approach. Figure 3.7 (a) shows the downstream portion of the approach (i.e. the portion closest to the intersection) and Figure 3.7 (b) shows the upstream portion of the same approach. In this example the viewing angles are not from the same direction, that is, the upstream is viewed from the South end while the downstream is viewed from the North end. The top right corner in Figure 3.7 (a) and the top left corner in Figure 3.7 (b) constitute the region of overlap.



(a) (b)

Figure 3.7: Example Views using Two Cameras (a) View of Upstream Portion of Approach, and (b) Example of Downstream Portion of Approach.

To process the vehicles and obtain their trajectories, having synchronization between the videos from the two cameras is very critical. The first step towards having this synchronization is having synchronized clocks for the recording. Since there are two cameras recording the same vehicle, care was taken to make the clocks in the two cameras have the same time. To reduce file sizes and minimize the likelihood of data corruption and data loss, video is stored as a series of 10 minute clips. Each video clip is named as a combination of camera name and starting time stamp of that particular segment. So, a particular vehicle would appear in two videos with different names corresponding to the two cameras. Therefore the second step is to have synchronization between the two videos opened by the software by taking into account the videos' starting time stamp (which is obtained from the video name). Even though the software tries to synchronize the two opened videos automatically, sometimes there would be minor differences. The videos are finally synchronized by manually matching the position of a test vehicle in the two videos at an overlapping detection line, i.e., a set distance from the intersection captured in both videos.

Once the videos have been synchronized, the ExtractData software component allows the analyst to lock the two videos together to step forward/backward synchronously when being reviewed by an analyst. That is, if one video is forwarded by the analyst, the other video is automatically forwarded by the same number of frames. But in some cases, the complete vehicle trajectory may not appear in the same pair of videos opened. A part of it (from a single camera) may be in one video segment and the remaining trajectory needs opening of the successive video. Provision is made to maintain such synchronization through the transition from one video clip to the next.

3.2.6 PET Data Extraction

The video analysis software is also used to evaluate the Post Encroachment Time (PET). PET refers to the time lapse between the end of encroachment of a turning vehicle and the time when the through vehicle enters the potential area of collision. PET is also used as a surrogate measure of safety. Any change in the behavior of through vehicle drivers resulting from the treatments can be expected to be reflected in the improved safety of interactions between these vehicles and left-turn vehicles, which is in-turn expected to be captured as an increase in the PET. Since there are two through lanes, there are two areas of conflict. These two areas are marked on the videos using the SaveGrid software component prior to starting the analysis, Figure 9. Next, using the ExtractData software component the analyst extracts the time stamps of the end of encroachment of left turning vehicle and the time of arrival of the through vehicle at the area of conflict. The difference between these timestamps is the PET.

Figure 3.8: Screenshot of Software Setup for Post Encroachment Time Data Extraction

3.3 DATA SAMPLING

For each mainline approach to an intersection a minimum of one week of video was recorded for each analysis time period, i.e., before and after treatment installation. This created a data set representative of each day of the week. Each day approximately 16 hours of video was collected during the daylight (and twilight) hours, resulting in ninety-six (96) 10-minute videos. Initial data reduction tests showed that each 10-minute video clip required about 4 hours (potentially more depending on the traffic flows) using the developed software. As this reduction process is highly resource intensive it was not possible to extract data from all video clips within a reasonable time period. Thus, a video sampling plan was adopted to capture a cross section of the recorded video. A 10 minute sample is extracted from each hour as representative data for that hour. The ten minute period selected for each consecutive hour is shifted by ten minutes in an attempt to avoid a data collection bias over the day. For example, if the first 10 minute video selected for a given day has a start time of 6:50 AM, the next video selected for the same day would be 7:40 AM, then 8:30 AM and so on. Similarly for the next day, the first video selected has a timestamp of 6:40 AM, then 7:30AM and so on.

Though the objective of this phase of research is to develop a methodology to obtain the surrogate measures considered, the developed methodology has been put to test through an example case of evaluating the effect of a safety treatment at the study intersection. Sampling of vehicles was also done in this context. Two of the surrogate measures being considered in this study are acceleration deceleration profile of through vehicles, and their intersection entering speed. According to the methodology applied in this effort, these surrogate measures are obtained from the time position data of the vehicles. Since the treatment is expected to improve the safety of interactions between left-turn vehicles and opposing through vehicles, the obtained surrogates are also chosen to capture these interactions. It follows that it is pertinent to capture the speed, and acceleration deceleration profiles of the through vehicles only in presence of left-turn vehicles. Therefore, data is extracted (according to the above sampling plan) for only those through vehicles satisfying the following conditions:

- There is an opposing left-turning vehicle that crosses the intersection while the through vehicle is within the approach area under study.
- The opposing left and through vehicle are both facing a green signal indication, thus the opposing left should only proceed if it has a sufficient gap.
- There is no standing queue of through vehicles at the intersection (such as the tail of a queue formed during the red phase prior to the green) which would affect the behavior of the approaching through vehicle.

These conditions are set as they capture the behavior of vehicles directly related to the objective of this effort, i.e. measuring the effect of the installed treatments on the vehicle behavior during potential conflicting intersection movements. By extracting data exclusively for these vehicles during the remaining 10 minute clips the data extraction time per 10 minute clip is reduced, allowing for more vehicles of interest to be sampled. Moreover, to obtain a baseline dataset of the behavior of vehicles approaching an intersection in the absence of opposing left turns, the time position data is extracted for some of the through vehicles approaching the intersection uninterrupted, for a representative day. Data from these vehicles allows for a determination of driver behavior in the absence of left turning vehicles.

3.4 DATA QUALITY ISSUES

As already discussed, the 900 ft approach zone is covered by two cameras. The viewing angles are relatively flat, on the order of 5 degrees. Thus, as the distance from the camera increases the detection lines in the perspective view appear closer, i.e., fewer pixels between lines. Consequently, the difficulty in identifying the frame in which a vehicle touches the detection line increases. The formula for calculating the observed speed based on the frame numbers recorded by the user is as follows:

$$V = (d*f/F)/1.4667 \qquad (3.1)$$

where:

f: frames per second in the video

d: inter-detector spacing

F: number of frames recorded by user for the vehicle to travel between two detector lines

V: observed speed in mph

If the user makes an error in recording the correct frame, the recorded speed is either greater than or lesser than the actual speed of vehicle depending on the nature of the error. The relationship between the actual and the recorded speed is given by the following formula:

$$V_{n+k} = f*d*v/(f*d + k*v*1.4667) \quad (3.2)$$

$$V_{n-k} = f*d*v/(f*d - k*v*1.4667) \quad (3.3)$$

where:

f: frames per second in the video

d: inter-detector spacing in feet

v: actual speed of vehicle in mph

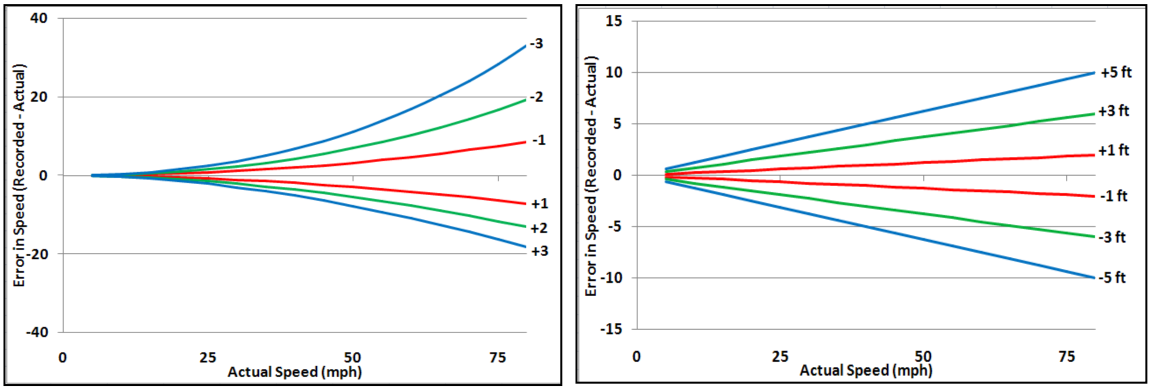
k: error in the number of frames recorded

n: correct number of frames in which the vehicle travels d distance

V_{n+k}: Recorded speed when k more frames recorded

V_{n-k}: Recorded speed when k less frames recorded

It is seen that the speed measurement granularity error is dependent on both the distance between the two detection lines for which the speed is being calculated and the vehicle speed. In the current study, the distance between consecutive detection lines is 40ft. The differences in the recorded and actual speeds corresponding to an error in frame recognition at 40 ft inter-detector spacing, using a video with a frame rate of 30 frames per second is shown in Figure 3.9 (a) while the difference corresponding to an error in the calibrated distance is shown in Figure 3.9 (b). The number beside each curve represents the error in recognizing the correct frame or the error in calibrated distance.



(a) **(b)**

Figure 3.9: (a) Error in speed corresponding to error in frame recognition (b) Error in speed corresponding to error in distance calibration

It can be seen in Figure 3.9 that while the difference in speed varies linearly with the error in calibrated distance, it varies non-linearly with the error in frame recognition. Moreover, at the same actual speed, negative values of "*k*" return greater errors in speed than those provided by positive values of "*k*". A basic explanation for this anomaly is that the error in frame recognition is in the denominator in the expression for recorded speed and that increasing and decreasing the denominator, though by the same quantity, will have different proportional effect on the result of the expression. This can be understood clearer by a closer look at the expression for calculating the difference between the recorded speed and actual speed corresponding to error in frame recognition.

$$V_{n+k}-v = f*d*v/(f*d + k*v*1.4667)-v$$

$$= -k*v^2*1.4667/(f*d + k*v*1.4667) \quad (3.4)$$

$$V_{n-k}-v = f*d*v/(f*d + k*v*1.4667)-v$$

$$= k*v^2*1.4667/(f*d - k*v*1.4667) \quad (3.5)$$

These expressions show that the difference in recorded and actual speeds is proportional to k/(C±k), where C = *f*d,* which is non-linear. They also show that the difference between recorded and actual speed is dependent on whether the error in frame recognition is positive or negative.

In addition to the possibility of errors in the collected data, the process of generating speed from video at a frame resolution of 30 frames per second gives discrete speed readings which results in noise in the data. As seen below in the Figure 3.10 (a), irregularities exist in the speed plot of the raw data. These irregularities are a result of data collection methodology, where discrete speed values are determined based on the number of frames required for a vehicle to travel the fixed distance between two detector lines. (In this study, the distance between consecutive detection lines was 40ft.) Some irregularities are also a result of inherent error in identifying the 1/30 of a second frame in which a vehicle crosses a detector line, as already discussed as potential sources of error. As the distance from the camera increases, the potential for this type of error increases due to the perspective view.

It needs to be seen that the excursions in the raw data at each detection line are not independent of each other. The sum of all the excursions should sum up to zero. To illustrate this concept, let us consider a small example. Let us assume that it takes x frames each for a vehicle to travel between detection lines 1 and 2, and 2 and 3 respectively. If the user incorrectly identifies detection line 2 by saving the number of frames it took the vehicle to travel between lines 1 and 2 as “x-1” frames, but correctly identifies detection line 3 by saving the number of frames as “x+1”, the second excursion (saving “x+1” instead of actual “x” frames) is a result of the first excursion, and that the sum of excursions is zero. This dependency is what makes the raw data plot not as noisy as it looks. This dependency is also what makes a moving average smoothing of the raw data reduce the noise at an order of n (much faster) instead of $\sqrt{n}$, if the excursions were independent.

The above analysis showed that there is a possibility of errors creeping into the collected data and explains the potential sources and magnitude of this error. Given these sources of noise in the data, some method is required to smooth the data and then validate the obtained profiles. Hence, low pass filters were developed to assess the accuracy of data collected and apply any corrective steps, and section 3.8 presents a deeper discussion on this.

3.5 SMOOTHING ALGORITHM

To verify the accuracy of the data collection methodology and to obtain the optimal smoothing algorithm, geographic positioning system (GPS) probe data was used. Use of geographic positioning system (GPS) equipped probe vehicles to collect comparison data for validation is a common practice. The first set of comparisons is performed between the data from GPS probe vehicle runs on the study site and the speeds generated using the software. The GPS equipment provides second-by-second location (latitude and longitude) data of the probe vehicle for several validation runs. A time position data, and from it an acceleration deceleration profile of the vehicle are generated with this data. The same vehicle is identified in the video and processed in the software to extract its position and acceleration deceleration profiles. The profiles from the two methodologies are compared to validate the accuracy of the data collection methodology being used in the study (Figure 3.10).

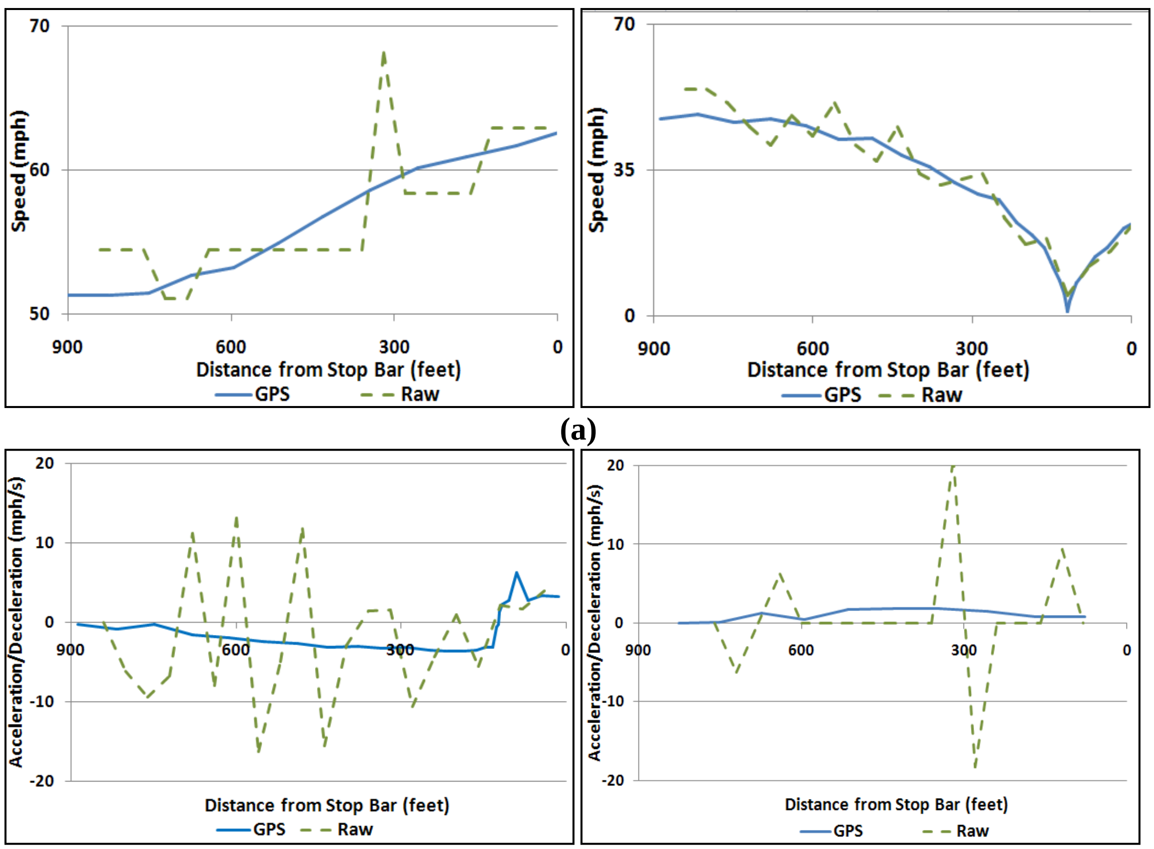


Figure 3.10: Example graphs showing (a) vehicle speed and (b) acceleration/deceleration profiles using custom software and comparing them with GPS data.

A smoothing algorithm should be chosen such that it removes the nominal irregularities in the raw data due to discretization of the speed values, but does not smooth out the higher values of accelerations or decelerations of the vehicles that occured. To obtain the algorithm that satisfies this requirement, a heuristic approach has been adopted. Various smoothing algorithms have been applied on the speed and acceleration data obtained. The simplest algorithm consists of an un-weighted moving average, replacing each point in the data with the average of 'm' adjacent points where m is a positive integer called the smooth width. For example, for a 3-point smooth,

$$S_j = (Y_{j-1} + Y_j + Y_{j+1})/3 \tag{3.6}$$

where S_j is the j^{th} point of the smoothed data, Y_{j-1}, Y_j and Y_{j+1} are the $j\text{-}1^{th}$, j^{th} and $j\text{+}1^{th}$ data points before smoothing. Three-point, five-point and seven-point moving average algorithms have been tested on the data. In addition, weighted average smoothing functions are also tested. A weighted average is any average that has multiplying factors to give different weights to different data points. For example, if we are taking the average of the values of Y_4 for raw data, 3-point moving average, and 5-point moving average, which we call as "1+3+5" weighted average smoothing; we obtain the expression for S_4 as

$$\begin{aligned} S_4\,(1+3+5) &= (Y_4 + S_4(\text{3-point moving average}) + S_4(\text{5-point moving average}))/3 \\ &= (3Y_2+8Y_3+23Y_4+8Y_5+3Y_6)/45 \end{aligned} \tag{3.7}$$

where S_j is the smoothed j^{th} point, Yj-2, Y_{j-1}, Y_j, Y_{j+1} and Y_{j+2} are the $j\text{-}2^{th}$, $j\text{-}1^{th}$, j^{th} $j+1^{th}$ and $j+2^{th}$ data points before smoothing.

So, we can clearly see that we assign different weights to different raw data points for finding the weighted average. This approach gives the highest weight to the central value and the weight decreases as we move farther. The following smoothing algorithms are tested on the data.

- Three-point moving average
- Five-point moving average
- Seven-point moving average
- 3+5 weighted average
- 3+5+7 weighted average
- 1+3+5+7 weighted average

Figure 3.11 shows the results of applying various smoothing algorithms mentioned above to the raw speed data and comparing them with the speed profile obtained from GPS data for a sample vehicle run. Similarly, Figure 3.12 shows the results of applying various smoothing algorithms to the raw acceleration/deceleration profile and comparing them with the profile obtained from GPS data.

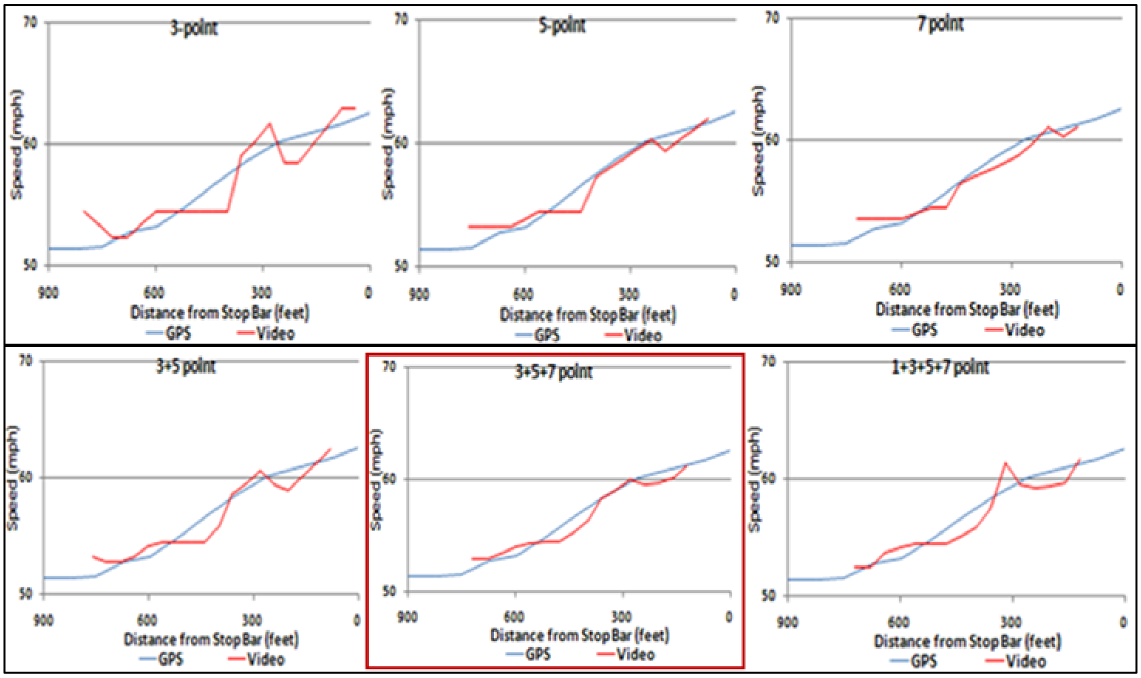


Figure 3.11: Plots showing the vehicle speed profiles after applying various smoothing algorithms on the raw data

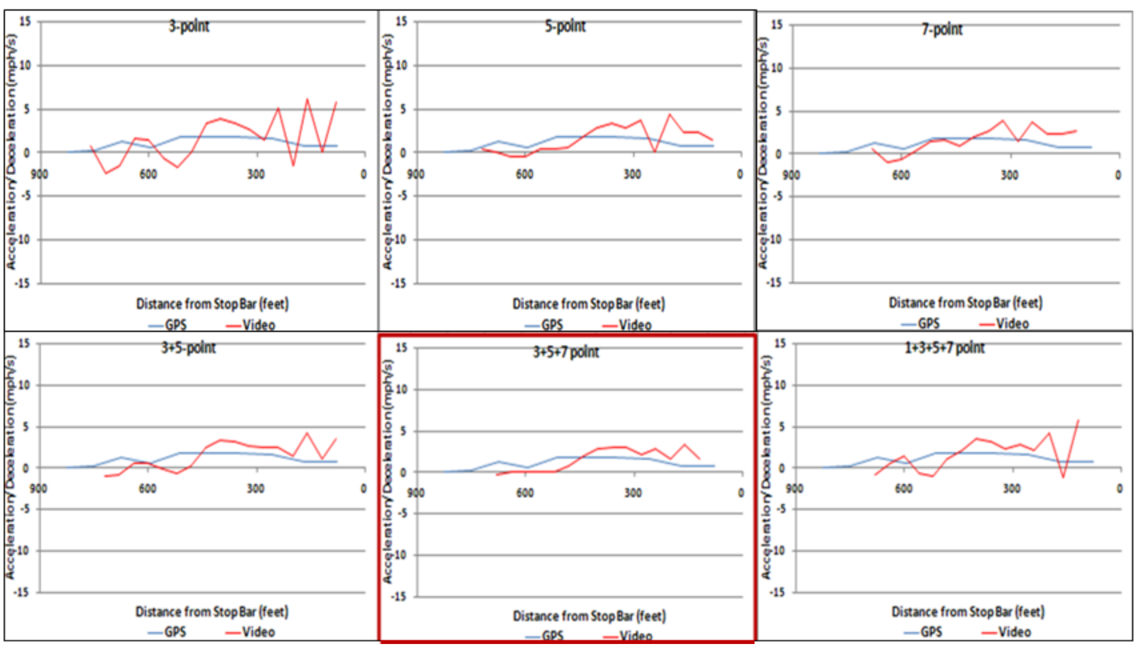


Figure 3.12: Plots showing the vehicle acceleration/deceleration profiles after applying various smoothing algorithms on the raw data

Visual interpretation of the Figures shown above indicate that “3+5+7” smoothing algorithm is the optimal filter which produces speed and acceleration-deceleration profiles closest to the GPS data. The graphs inside the red outline in both the Figures 3.11 and 3.12 show the results of visual interpretation. In addition to visual interpretation, it was also decided to compute mean squared error for the speed and acceleration-deceleration data obtained from video and GPS. Assuming GPS data as the base data, the smoothing algorithm which gives the least mean squared error taken as the optimal (of those tested) smoothing algorithm. Table 3.1 shows the results of the mean squared error computation.

Table 3.1: MSE for smoothing algorithms assuming GPS data as the ground truth

Source	Speed MSE				Acceleration/Deceleration MSE			
	Run1	Run2	Run3	Run4	Run1	Run2	Run3	Run4
Raw Data	4.2064	223.70	6.1851	7.7086	34.628	221.15	54.943	59.573
3-point	2.3092	2.7471	2.0545	2.3736	3.2787	14.240	4.3531	6.8987
5-point	2.1231	1.2894	1.4883	0.7152	1.1828	4.1645	2.1921	2.3559
3+5 point	2.1249	1.4812	1.496	1.0661	0.9090	4.0613	1.0823	2.8395
7-point	2.0728	1.5019	1.3690	0.4884	1.1237	3.5335	1.2424	2.1304
3+5+7 point	2.0293	1.3009	1.115	0.5345	0.8272	1.6524	0.5876	1.8850
1+3+5+7 point	2.1102	1.7117	1.3819	1.1662	1.9950	3.4722	3.5620	5.7996

It shows four runs where each run represents one run of the GPS probe vehicle through the 900 feet approach to the intersection. It can be seen from the Figure that for the acceleration-deceleration data, “3+5+7” smoothing gives the least MSE for all four runs while for speed, “3+5+7” gives the least MSE for two runs while it gives the second least MSE for the other two runs. Overall, it can be concluded that “3+5+7” smoothing algorithm is the optimal smoothing algorithm. This smoothing is defined as:

$$S_4\ (3+5+7) = (S_4(3\text{-point moving average})+S_4(5\text{-point moving average})+S_4(7\text{-point moving average}))/3$$

$$= (15Y_1+36Y_2+71(Y_3+ Y_4+ Y_5) +36\ Y_6+15Y_7)/315 \qquad (3.8)$$

where:

S_j: smoothed j^{th} point

Y_{j-3}: $j-3^{th}$ data point before smoothing

Y_{j-2}: $j-2^{th}$ data point before smoothing

Y_{j-1}: $j-1^{th}$ data point before smoothing

Y_j: j^{th} data point before smoothing

Y_{j+1}: $j+1^{th}$ data point before smoothing

Y_{j+2}: $j+2^{th}$ data point before smoothing

Y_{j+3}: $j+3^{th}$ data point before smoothing

The speed and acceleration-deceleration data of the sampled vehicles obtained from the video reduction are smoothed using the “3+5+7” filter before proceeding to the comparison of before and after treatment data to evaluate its effectiveness. Some examples of the speed and acceleration deceleration profiles after applying the “3+5+7” smoothing algorithm are shown in Figure 3.13 and Figure 3.14 respectively.

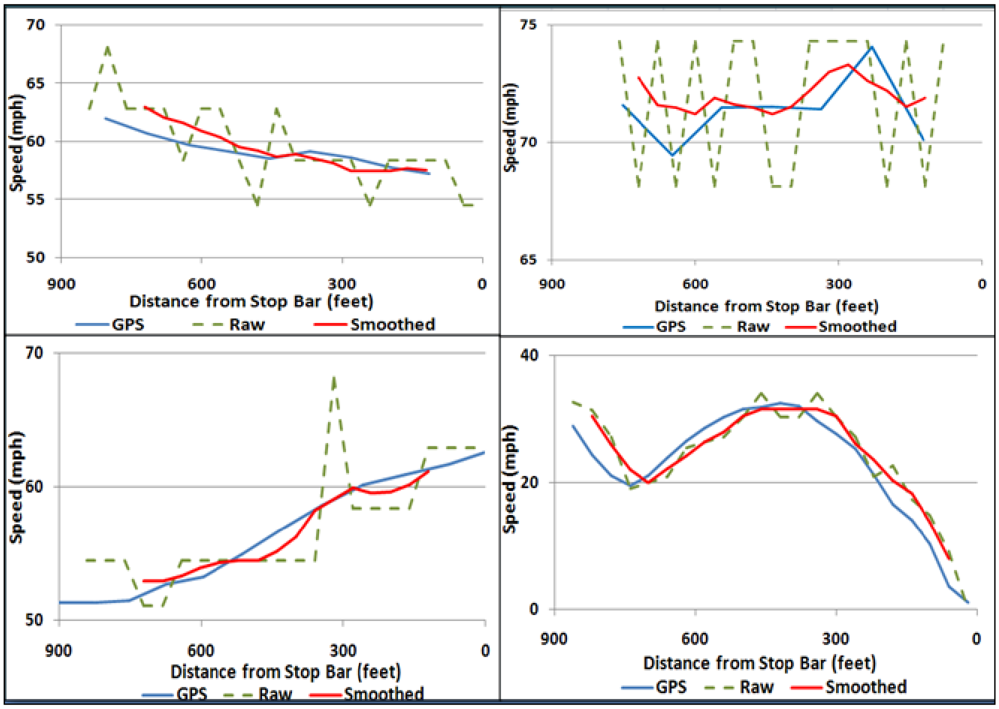


Figure 3.13 Effect of "3+5+7" weighted average smoothing algorithm on the raw speed profiles

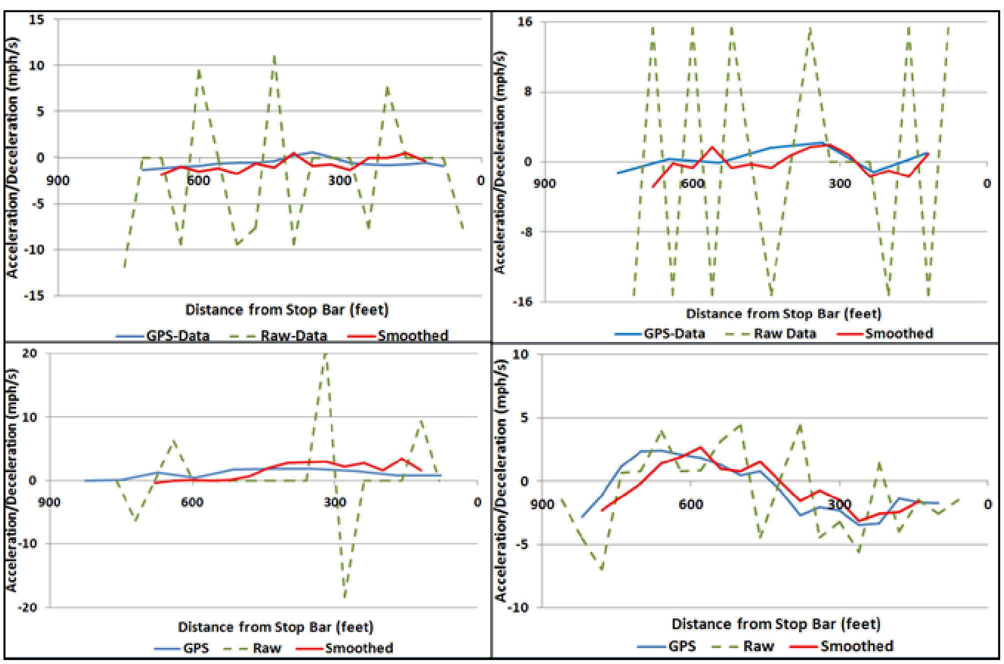


Figure 3.14 Effect of "3+5+7" weighted average smoothing algorithm on the raw acceleration/deceleration profiles

3.6 EVALUATION OF THE TREATMENT

To evaluate the effectiveness of this data collection methodology, its ability to produce data sets for several surrogate safety measures at the study intersection was evaluated. These surrogates were considered to be potentially useful for comparing the before and after treatment data and were evaluated for the southbound and northbound approaches of the study intersection. The treatment was expected to bring a change in the driving behavior of the through vehicle drivers as they approach the intersection. This change is expected to reflect on the interactions between these through vehicles and the corresponding opposing left-turning vehicles, thereby increasing safety of their interactions. The behavior of the through vehicles and safety of their interactions with opposing left-turn vehicles are captured in three quantitative measures:

a) Acceleration/deceleration profile of through vehicles,
b) Post Encroachment Time (PET), and
c) Intersection entering speed of through vehicles.

Changes between the before and after observations, if any, may be at least partially attributed to the implemented safety treatment and likely indicative of the potential treatment impacts. The objective of this phase of research however was to develop the data collection methodology and understand its strengths and weaknesses. The videos collected at the study intersection were processed to obtain the speed and acceleration-deceleration profiles of both northbound and southbound through vehicles using the custom software described earlier.

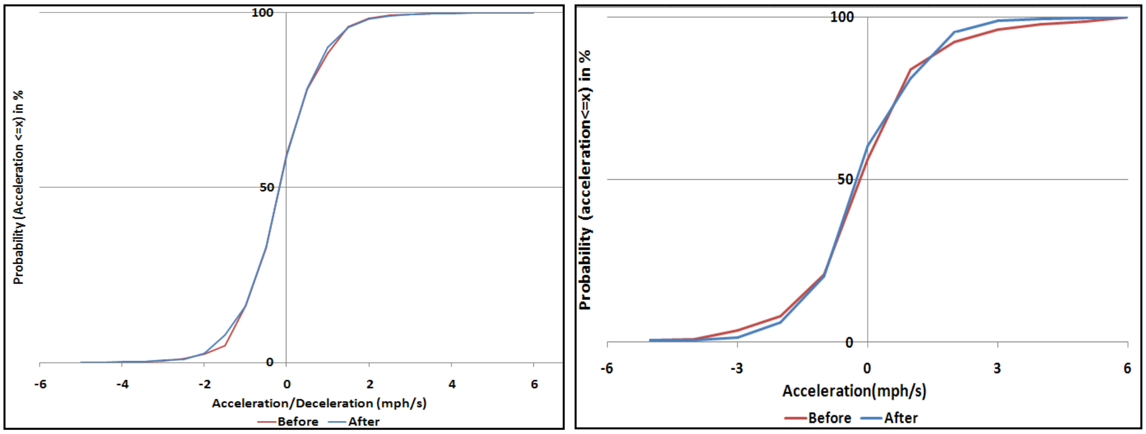
For this analysis it is recalled that speed and acceleration/deceleration profiles for through vehicles are measured from the boundary of the intersection proper (defined as the stop bar) to a position approximately 900 ft upstream. Raw data collection consists of time versus position data collected over the data collection zone. The data collection zone is divided into 40 ft intervals with the time a vehicle crosses each 40ft marker determined through the use of video recordings. Thus, over the 900 ft data collection area, speed and acceleration/deceleration trajectory data are collected, containing approximately 25 discrete data points at 40 ft spacing.

It is noted that for the given intersection, traffic demands, and data collection periods, conflict opportunities (i.e. the arrival of a through vehicle with no standing queue on the through approach and a vehicle turning left at the intersection) were greater on the southbound approach. In the following analysis before and after southbound results are based on 300 and 297 through vehicles, respectively, while the northbound before and after results are based 42 and 44 through vehicles, respectively (counts are obtained from the sampled video data for processing). These are the number of vehicles that were processed from the sampled video clips and ones that satisfied the conditions mentioned in section 3.3 to choose vehicles to process.

3.6.1 Vehicle Acceleration and Deceleration

First surrogates considered were two complementary analyzes of the acceleration/deceleration profiles of through vehicles. The first acceleration/deceleration analysis used data from the entire vehicle trajectory as the vehicles approach the intersection. This would provide an overall trend of any shift in the distribution of acceleration deceleration values, if any due to the treatment. But a conflict may also be characterized by sudden application of brakes to evade a potential collision, as perceived by the driver. This behavior is expected to be reflected in the maximum deceleration rate experienced by the vehicle in its trajectory, and is considered to be the complementary measure analyzed. A significant difference in the before and after treatment distribution of maximum decelerations may indicate a change in the number or severity of conflicts. Previous studies have also shown the application of such a measure in surrogate safety analyses, which is called “deceleration rate”.

Figures 3.15 (a) and 3.15 (b), show the cumulative distribution function (CDF) of the observed acceleration values for the southbound approach (300 vehicles), and northbound approach (44 vehicles) respectively. The southbound CDF plot shows that the probability of having decelerations in the range of -1 to -2 mph/s is marginally higher in the after treatment data than in the before data. Other than this difference, the two curves overlap quite consistently, implying that the difference in the distributions of accelerations and decelerations in the before and after southbound data is very minor, if any. The CDF plot of the northbound before and after data indicates a slight decrease in the likelihood of decelerations rates in the -1 to -4 mph/s range in the after data. However, the likelihood of decelerations greater than -4 mph/s is same in both the before and after data. In examining Figure 3.15, it is interesting to note that the deceleration values fall within the comfortable deceleration value of 7.63 mph/s (11.2 ft/s/s) typically utilized in intersection signal design (AASHTO 2004).



(a) (b)

Figure 3.15: CDF of the acceleration/deceleration profiles (a) southbound approach (b) northbound approach

Figures 3.16 (a) and 3.16 (b), show the cumulative distribution function (CDF) of the observed maximum deceleration values for the southbound approach (300 vehicles), and northbound approach (44 vehicles) respectively. The CDF plots of maximum decelerations for the southbound approach shows a potentially slight increase in the recorded maximum deceleration in the after treatment data collection, however, the magnitude is likely insignificant. There is a minimal difference in the range of -1 mph/s to -2 mph/s. However, there is no perceivable difference in the CDF of the deceleration values greater than -3mph/s. The northbound maximum deceleration data shows higher frequency for the before treatment data in the range of -3 mph/s to -5 mph/s but the after data has higher frequency for maximum decelerations greater than -5 mph/s. There are again differences in the lower range deceleration rates; however, the trend is opposite that seen on the southbound approach. Thus, similar to the aggregate acceleration/deceleration data the maximum deceleration data contains no significant changes that would indicate a difference in before and after treatment conflicts or safety performance.

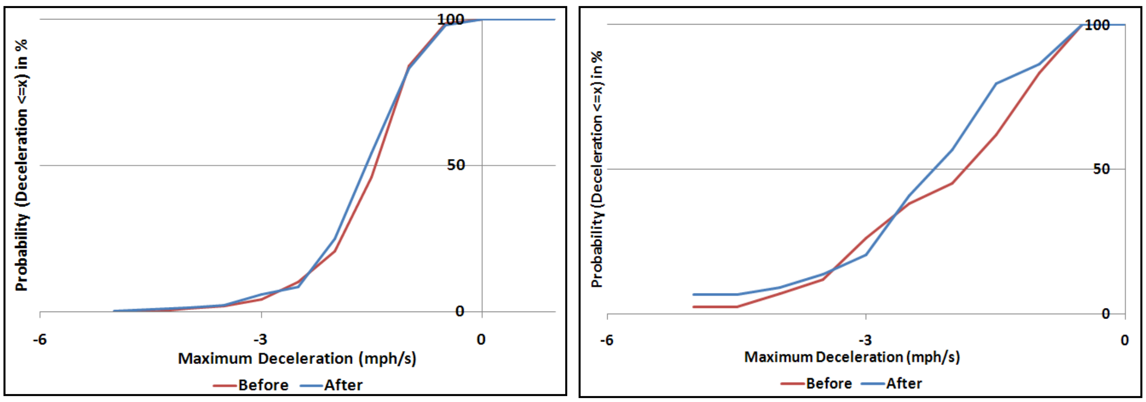
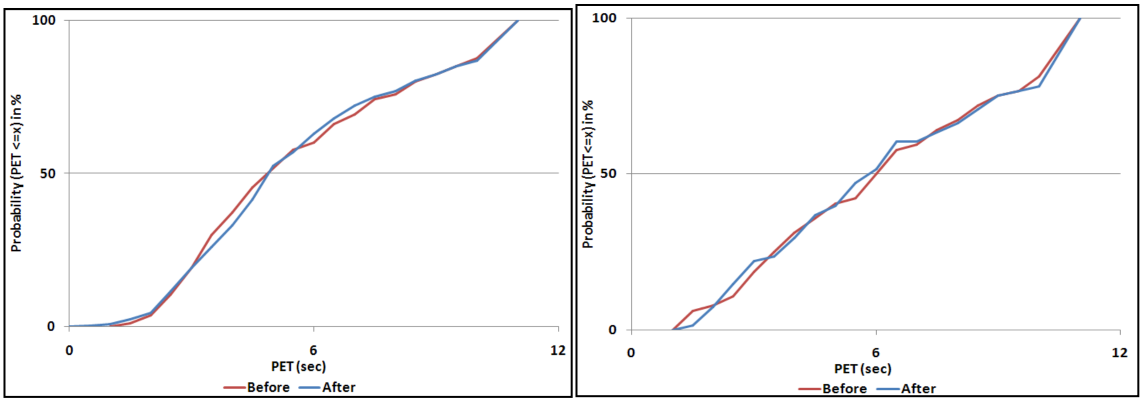


Figure 3.16: CDF of the maximum decelerations (a) southbound approach (b) northbound approach

Both complementary analyzes of the "acceleration deceleration profile" surrogate measure illustrate the ability of the methodology to characterize large numbers of vehicles not only in terms of speed but also of deceleration distributions.

3.6.2 Post Encroachment Time

Post Encroachment Time (PET) is the next surrogate measure considered in this analysis. PET is the time lapse between the end of encroachment of the turning vehicle and the time that the through vehicle arrives at the potential point of collision. Any increase in the level of safety in the interactions between left-turning vehicles and opposing through vehicles resulting from the treatments can be expected to be reflected in an increase in the PET. Figures 3.17 (a) and 3.17 (b) show the CDF plots of the PET values between the left turning vehicles and the through vehicles for the southbound approach and northbound approach respectively. Figure 3.17 (a) shows a small shift in the PET values from the range 4-5 sec to the range 5-6 seconds in the southbound direction. There is no appreciable change from the before to after treatment periods in the distribution of PET values greater than 6 seconds. The northbound data shows the reverse trend, with an increase in the number of PET values in the lower range (3-4 seconds) in the after data. Again, there is no notable trend in the higher PET values. The northbound PET CDF demonstrates no consistent shifting between the before and after PET behavior. As with the acceleration/deceleration observations the observed PET differences before and after treatment are all minor. Literature suggests that critical PET values fall within the range of 3 to 5 seconds. While some differences between the before and after treatment data are witnessed in this range there is no consistent pattern between the northbound and southbound approaches.

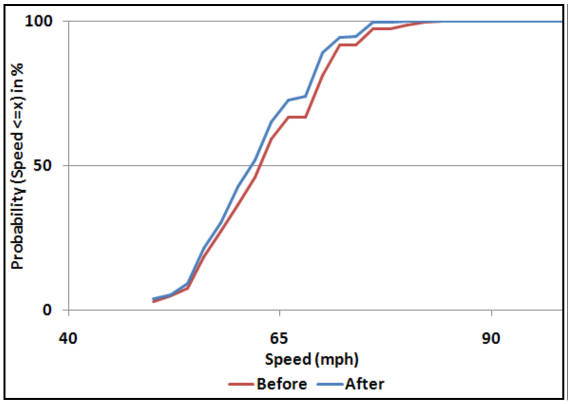


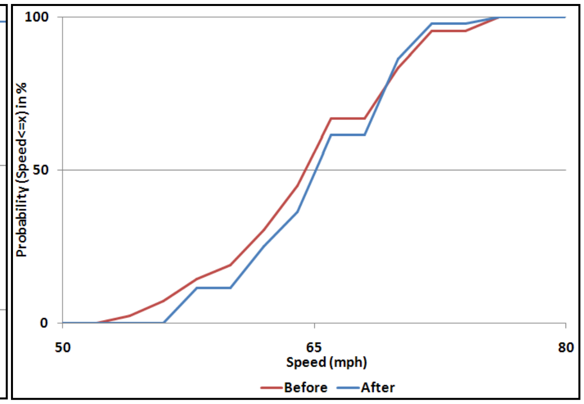
(a) **(b)**

Figure 3.17: CDF of the PET values for the before and after data (a) southbound approach (b) northbound approach

3.6.3 Through Vehicle Speeds

Figures 3.18 (a) and 3.18 (b) show the southbound approach and northbound approach before and after treatment speed distributions respectively for through vehicles as they enter the intersection proper. In the CDF plot corresponding to the southbound approach, some shifting of the speed distribution to lower values, on the order of 3 mph, is seen in the after treatment. The northbound data shows generally the opposite trend with a decreased likelihood of lower speeds after the treatment installation. So, both the approaches do not show the same trend in the variation of speed distributions.





(a) **(b)**

Figure 3.18: Speed of through vehicles entering intersection proper (a) southbound approach (b) northbound approach

Overall it is seen that all three surrogates considered in the study (acceleration/deceleration values, PET values, and intersection entering speed of through vehicle) show only minor differences in the distributions of the before and after data. The differences, if any, are small and often oscillating around zero, likely indicative of expected minor variability resulting from the data collection procedure and underlying randomness in the driver's behavior rather than a systemic treatment effect.

It may be possible that the treatment is subtle and did not have considerable effect on the safety of the interactions between the through vehicles and left-turn vehicles. The after-data at for the northbound approach were collected during the 3rd and 4th weeks after the treatment was applied and the after-data for southbound approach was collected during the 5th and 6th weeks after the treatment was applied. It might also be possible that the treatment had an immediate effect on the behavior of drivers but they might have adapted to the treatment nullifying its effect by the time data was collected. Secondly, one of the objectives of the research is to evaluate the effect of the treatment using surrogate measures for of crash data. Chapter 2 shows us that previous studies have doubted the effectiveness of surrogates as the collected data often contains a mixture of serious and non-serious conflicts. This might also be one of the potential reason for the considered surrogates not showing any effect of the treatment. It is unclear, as of now, how to recognize serious conflicts among all conflicts for the considered surrogates. With respect to the current study, literature does not provide information on threshold values for the surrogates considered in this study and this might be one of the reasons for the analysis not showing effect of the treatment, if any. Nevertheless, the data collection technique has been shown to be effective at producing the data of both profile-based and point-based surrogate measures.

3.7 SUMMARY

This phase of research has illustrated a methodology to collect both profile based and point based surrogate measures. A review of literature (chapter 2) has shown that on field data collection of profile based surrogate measures such as speed profile, acceleration deceleration profile, TIT and TET has not been attempted much due to the difficulty in collecting such data. Such measures have been collected either through simulations or using GPS equipped vehicle data. The methodology developed in this phase of research successfully extracts the surrogate measures - speed, acceleration/deceleration profiles, and PET from video data. The video recording and post processing adopted in this research enables field data collection of profile-based surrogate measures. As an example case, this methodology has been applied to evaluate the effectiveness of a safety treatment at an intersection by comparing the before and after treatment surrogate data.

A semi-automatic method using custom software was developed for extracting the surrogate safety data from videos as the methodology requires collecting speed data at fixed intervals along a longer stretch of approach road. This semi-automatic approach allows for the use of lower camera angles with larger perspective views thereby limiting equipment needs, a limitation of most of the automatic video detection equipment based approaches. The methodology involves extracting the trajectories (position-time data) of the through vehicles satisfying certain conditions, and extracting the speed profiles and acceleration-deceleration profiles from the trajectory data. Though this methodology has certain advantages, the cons of such an on-field profile-based surrogate data collection methodology could be understood.

There were a lot of quantities that needed to be determined before the data was collected. Profile-based surrogates would require collecting the surrogate measure over a certain length of approach road. Depending on the length being considered, it was first necessary to determine the number of cameras that would be required to cover the roadway. A camera test was conducted prior to the actual field deployment to determine some other key factors. First of these factors is the angle of view of the camera. If the camera is placed at a high angle, visibility would be very good as the camera would directly look down with less perspective angle, but the viewing area would be less. If the angle were less, the camera would cover larger area but would have perspective in the view. The second factor is the height at which the camera needs to be placed. This test demonstrated the optimal height for placing the camera to obtain the best balance between visibility and area of coverage.

There were also issues that were encountered with the portable system during the data collection period. For example, inclement weather conditions were not conducive for data collection. There were long periods of cloudiness which hampered the charging of the batteries from the solar cells and gave rise to power related issues. Placing of trailers away from the camera in a sunny place led to greater cable length, loss of data, and damage to adapters. Grade of the approach road gave rise to problems with camera viewing area, especially for the northbound approach. There were intermittent issues with respect to network connectivity, movement of camera out of the view due to wind etc.

The collected data has two major sources of potential error. The first source of error comes from the user not recording the correct frame when the vehicle touches the detection line. The second source of error comes in the consideration that the distance between every pair of detection line is 40 ft. Even though cones were placed by an on-field measurement of 40ft, some errors might still be there in these measurements. These errors combined with the fact that the frame rate of the video recording of 30 frames per second leads to discrete speed data points and creates noise in the raw data. This required the development of low pass filters to smooth the data, and it was found that a “3+5+7” weighted average algorithm was optimal to smooth the data. This experience shows that such studies would be prone to noise in data and would require developing smoothing algorithms.

Even with these limitations, the methodology successfully extracted the surrogate measures planned to be collected for this phase of research. However, for neither the northbound or southbound approach data do the surrogate measures considered show that the treatment has any significant effect on the safety of interactions between left-turning vehicles and opposing through vehicles. All three surrogates considered in the study (acceleration/deceleration values, PET values, and through vehicle intersection speeds) show only minor differences in the distributions of the before and after data. The differences, if any, are small and often statistically insignificant (oscillating around zero), likely indicative of expected minor variability resulting from the data collection procedure and underlying randomness in the driver's behavior rather than a systemic treatment effect.

From this study it is not clear if the surrogates considered did not show any significant difference before and after the application of the treatment because the treatment is so subtle that it did not have any discernable effect on the interactions between left-turn vehicle and opposing through vehicle, or if the considered surrogates are not effective in capturing these interactions. It is also possible that the sampled data is insufficient to capture any difference in the before and after treatment data, if any present. Chapter 2 also shows us that previous studies have doubted the effectiveness of surrogates as the collected data often contains a mixture of serious and non-serious conflicts. This might also be one of the potential reasons for the considered surrogates not showing any effect of the treatment. It is unclear, as of now, how to recognize serious conflicts among all conflicts for the considered surrogates.

Hence for the next phase of the research, it was decided to evaluate the effectiveness of these surrogates by collecting surrogate data at more intersections having high, medium and low crash frequencies. But, this phase of this research showed that collecting acceleration deceleration profiles is equipment and labor intensive task. The methodology presented in this research required a portable data collection system that needed to be left on field and monitored remotely. Even though this methodology had certain advantages, various issues with the process were discussed in this chapter. As such, a replication of this study to capture the same surrogate measures for multiple locations is likely time and resource prohibitive for most circumstances. Collection of speed profile data also suffers from these limitations. Unless automated video post-processing systems are developed to quicken the process of profile-based surrogate data collection, such data collection process might not be scalable. Spot speed can be a better surrogate in terms of labor and equipment requirements but a definite boundary value which differentiates a crash or non-crash event cannot be defined for it.

In lieu of these observations, the experience from the above mentioned studies and previous research works, it can be hypothesized that PET has a high likelihood of providing a usable and cost-effective surrogate measure. It is relatively easy to measure as it requires collecting only two timestamps for each PET data point and a PET value of zero differentiates crash and non-crash events. Therefore, the next phase of the research digs deeper into the potential surrogate measure – PET, and focuses on evaluating its effectiveness and applicability. The interaction being studied is still between left-turn vehicle and opposing through vehicle.

CHAPTER 4: PHASE 2

4.1 INTRODUCTION

Chapter 3 described phase 1 of this research that developed a methodology to collect profile-based and point-based surrogate measures in the field. The methodology was used to collect three potential surrogates – acceleration-deceleration profile, PET, and through vehicle intersection entering speed. As an example case, this methodology was applied to evaluate the effect of a safety treatment at an intersection in rural Georgia. The experience from the first phase of research has shown that it is necessary to further evaluate the effectiveness of the considered surrogates. This further evaluation requires collecting data at additional locations. The first phase has also shown that collecting acceleration-deceleration profile data at multiple locations is not practical using the developed methodology as it is highly labor intensive and time consuming. PET on the other hand requires only two time stamps, a single camera and small area to view, thereby making PET amenable to collection at multiple locations, and importantly, a realistic candidate for part of a broader safety evaluation program at a state or local agency.

This phase of research delves deeper into the evaluation of effectiveness of PET as a surrogate measure for crashes. The conflict being studied is between a left-turn vehicle and an opposing through vehicle at a signalized intersection. This phase involves data collection at additional intersections, analysis of the properties and distribution of PET, and finally evaluating its effectiveness as a surrogate measure. This study, in addition to developing a cost-effective data collection procedure for obtaining a statistically sufficient PET data sample, is expected to support and increase confidence in the use of PET as an effective surrogate measure of safety. This phase of the research addresses objective (ii), described in chapter 1.

4.2 METHODOLOGY

This section will discuss the procedure adopted for collecting the data for this phase of research. It will be seen that this procedure is based on lessons learned from the efforts of phase 1, described in chapter 3. This section will describe a porTable and cost-effective equipment set-up developed for this phase of research. Significant effort was expended for selecting the study intersections and hence these steps are also described in this section.

4.2.1 Data collection

As with Phase 1 the primary data collection methodology is video recording of the traffic streams. The custom frame-by-frame video reduction software program developed for the first phase of research has been further adapted to increase the efficiency in extracting PET data. Details of these adaptations are presented in subsection 4.2.5 that describes modifications to the software in detail. In addition the data collection scheme was altered to allow for a minimum level of equipment deployment. One of the advantages the first phase of research had was that the cameras were placed at height of approximately 45 feet and had a view of the intersection with relatively low occlusions and a low perspective view ("low" for the intersection area but still had high perspective angle for the approach road). However, to accomplish this, permanent poles having a height greater than 50ft (on which the cameras were placed) were installed by GDOT. This is not a practical approach that could be replicated in a larger scale study involving numerous intersections, which is the objective of this research phase. Therefore, an effort was made to select locations with at least a section of elevated roadside and limit the equipment to standard tripods, able to extend to approximately 8 feet. Selecting locations adjacent to the study intersections that were higher than the intersection, allowed the camera to capture a view of the complete intersection and movement of vehicles. The cameras were also placed at such a location where they would not be conspicuous to the drivers and would not influence their driving behavior.

4.2.2 Crash Data

Accident records for the years 2006 through 2009 were processed to generate candidate intersections for the PET study. Data were analyzed from crash records from the crash database provided by GDOT and the *Critical Analysis Reporting Environment* (CARE) software (*23*), based on the GDOT crash data. CARE is an application that assists users in creating database queries for crash data, and obtaining information from publically available sources. State transportation agencies (DOTs) that use CARE provide the

Center for Advanced Public Safety (developers of CARE) the crash data for each year and then this data is converted into a format or dataset required by CARE that enables users of CARE to work with the crash data. CARE provides crash data analysis in the form of descriptive statistics, cross-tabulations, hot-spot determination, generating collision diagrams, and certain GIS capabilities.

At the outset, it is important to understand the various steps that the crash data analysis has gone though in this research. Details of the GDOT crash databases (formats that varied over the years, the various columns or data in the database) were discussed in chapter 3. It is generally known that crash data has limitations in terms of inaccuracies. There are crashes with missing RCLink data, missing milepost data and/or latitude/longitude information. Therefore all crashes at each intersection may not be identified due to this missing data. However, there were many other forms of problems that were encountered at various stages of the research. As mentioned in the paragraph above, the first source of crash data was from CARE software. The CARE software uses RCLink and milepost data from the datasets provided by GDOT. The initial efforts of analysis in this research used crash numbers from CARE software.

Chapter 5 discusses another phase of the research where many more intersections were considered for analysis. At this stage, the first problem with the crash data was identified. The issue was that some of the crashes identified by CARE database as belonging to an intersection were incorrectly attributed to that intersection. This is caused due to an error in identifying the correct RCLink for a crash while creating the initial database by

GDOT, and this error carried over to the CARE database as well. The details of the issue, its identification and resolution are explained in chapter 5. The crash numbers were accordingly corrected and the analysis was completed.

However, there were some additional issues with the crash data that were identified later. First of all it was found that the initial crash databases provided by GDOT had significant missing data in the months of November and December for the year 2008 (which obviously has an impact on the crash numbers used, as discussed in the previous paragraph). Moreover, a different format of crash database called “sanitized” databases were provide by GDOT for the study years which had intersection road names (major and cross road) instead of RCLink and milepost combination. From this new data, crash records for the study intersections were again filtered and verified for any data that were not obtained by from CARE software. Police reports for these crashes were again referred to for filtering out the records corresponding to opposing left-turn crashes, and also to identify any other errors/inaccuracies. It was found that additional crashes were identified using the sanitized databases when searched using road names. This meant that some of the crashes which missed out being assigned to the correct intersection due to error in assigning RCLink were identified from the sanitized database’s road names.

It was observed that each form of databases has its own limitations. The database that CARE software uses has some records wrongly assigned to intersections because of assigning a wrong RCLink. Missing information such as RCLink, milepost also causes some crashes not being assigned to an intersection. The sanitized database on the other hand has limitations in terms of problems with the road names. First of all, roads of an intersection can be represented in multiple ways. Sometimes they have road names and sometimes route numbers (For example, US Highway 20 or State Route 20). Even in these representations, the names can be abbreviated. This gives rise to various forms and combinations in which an intersection can be mentioned in the database. Lastly, there can be mistakes in the spellings of road names due to which some of the crash records can be missed in a query. Even with these limitations, more than 90% of crashes identified by CARE software were identified by road name search method. However a combination of both forms of databases can be used to obtain a better set of crash records for any intersection. Using this approach, crash numbers were recalculated for each intersection and the analysis was redone. Let us call this “final analysis”. The analysis in all chapters hereafter uses the final crash numbers obtained by a combination of both these methods.

4.2.3 Site Selection

It is expected that one of the major applications of PET will be to provide an initial evaluation of the effectiveness of any safety treatments or countermeasures applied at an intersection without having to wait the typical 3 year period for collecting the accident data. In such an application, for intersection improvements that do not impact capacity (e.g. new lanes), volume immediately before and after the treatment will typically be similar, with only minor volume fluctuations. Thus, the difference in the before and after PET data should be attribuTable primarily to the effect of the treatment, potentially allowing PET data to be an accurate surrogate measure of safety. However, in this effort it was not possible to collect before and after treatment data at sites where treatments were applied. Thus, to evaluate the effectiveness of PET as a surrogate measure for safety, it is advantageous to select intersection pairs which have similar operating conditions (AADTs, intersection control, lane configuration etc.) but different crash frequencies. In addition, the conflict between left-turning and opposing through vehicles is the primary PET conflict of concern in this effort. Therefore, the frequency of crashes which occurred due to this conflict is directly considered. First, the possibility of using functional classification as a basis for selecting pairs of intersections was evaluated on the assumption that a functional class would have roads with similar AADTs. But an analysis of AADT and functional class showed large AADT bandwidths for each functional class which means that there might be significant difference in the AADT values of two intersections having different crash frequencies. Moreover, to reduce potential confounding variables in the paired intersection PET analysis (such as different driver populations, etc.) a corridor level analysis of crash frequencies is used to select the candidate intersections pairs. Pairs of intersections on a corridor are selected having similar characteristics (AADT, lane configurations, and intersection control) but different crash frequencies. This facilitates a pairwise comparison to evaluate the effectiveness of PET.

An important aspect that needs to be mentioned here is that the AADT being considered is the major road AADT and not the total AADT. For selecting intersections, it was necessary to consider both crash frequencies and AADT together. Fortunately, the crash database obtained from GDOT had AADT values, but they are major road AADTs.

Hence only major road AADTs were considered in this selection process. For the initial data collection, three corridors (GA -8, GA-10 and GA-20) were selected in Gwinnett county, Georgia. Out of these, two intersections, GA 10 (Main St) at Henry Clover Blvd/Oak Rd and GA 10 (Main St) at Grayson Pkwy in Gwinnett county of Georgia were selected. The distance between the intersections is 1.2 miles, thus the intersections likely have similar population of drivers. The intersections have similar AADT counts, signal control, and geometries but have different crash frequencies between left-turn and opposing through vehicles, as summarized in Table 4.1. The crash frequencies are based on a RCLINK-based search using CARE.

Table 4.1: Characteristics of the intersections of GA 10 with Grayson Pkwy, and GA 10 with Henry Clower Blvd/Oak Rd

Intersection	4-year crash count	Major AADT	Type of Intersection	Number of Lanes (on major road)
GA 10 with Grayson Pkwy	29	41400	4-legged, signalized (protected-permitted left-turn)	2-through + 1 left-only
GA 10 with Henry Clower Blvd/Oak Rd	9	33630	4-legged, signalized (protected-permitted left-turn)	2-through + 1 left-only

Even though two such intersections were found, it was still difficult to select pairs of similar intersections at corridor level but having varied crash counts. Therefore, for the next pair of intersections, it was decided to relax the condition of being on the same corridor. Consequently, the second pair of intersections selected was that of Roswell Rd with Wieuca Rd, and Buford Hwy with Sugarloaf Pkwy. Characteristics of these intersections are summarized in Table 4.2.

Table 4.2: Characteristics of the intersections of Roswell Rd with Wieuca Rd, and Buford Hwy with Sugarloaf Pkwy

Intersection	4-year crash count	Major AADT	Type of Intersection	Number of Lanes (on major road)
Roswell Rd with Wieuca Rd	78	25594	4-legged, signalized (protected-permitted left-turn)	2-through + 1 left-only
Buford Hwy with Sugarloaf Pkwy	6	26210	4-legged, signalized (protected-permitted left-turn)	2-through + 1 left-only

4.2.4 Data Collection

Video data was collected primarily under day-light, non-inclement weather conditions. Initially, data was collected during off-peak hours from 10 AM to Noon and again from 1 PM to 4 PM at the first pair of selected intersections, GA 10 with Grayson Pkwy and GA 10 with Henry Clower Blvd/Oak Rd. The data was collected during off-peak hours because a visit to the site showed "signs" of the unavailability of sufficient gaps for the left-turn vehicles during peak hours, significantly limiting any left turns (and thus PET opportunities) during the permissive portion of the left turn phase and the incident data showed that approximately 70 percent of incidents at these intersections occurred during non-peak hours. Data collection at both intersections was carried out simultaneously to increase population consistency. Two days of non-peak hour data was collected at these two intersections. However, data collected at these two intersections indicated that the peak period as well as the non-peak should be considered, which would also give a better understanding of PET data distribution. So, at the first pair of study intersections, data was collected on three different days to allow for the inclusion of a peak period.

However, for the remaining intersections a time slot having a combination of peak and non-peak hours was used. Details of the different days and times data was collected at each intersection are presented in Table 4.3.

Table 4.3: Data collection dates and times at the study intersections

Intersection	Date	Time of Data Collection
GA 10 with Grayson Pkwy	4th October 2010	10 AM - Noon , 1 PM - 4 PM
	7th April 2011	10 AM - Noon , 1 PM - 4 PM
	6th May 2011	4 PM - 7 PM
GA 10 with Henry Clower Blvd/Oak Rd	4th October 2010	10 AM - Noon , 1 PM - 4 PM
	7th April 2011	10 AM - Noon , 1 PM - 4 PM
	6th May 2011	4 PM - 7 PM
Roswell Rd with Wieuca Rd	31st May 2011	7 AM to Noon
Buford Hwy with Sugarloaf Pkwy	3rd June 2011	2 PM to 7 PM

4.2.5 PET Data Collection

Video data was reduced in a laboratory environment using a modified version of the custom video reduction software discussed earlier. An example of the custom software interface is seen in the screenshot of the intersection of GA 10 at Grayson Pkwy, shown in the Figure 4.1. The red and blue lines and the blue numbers on the screen represent identifiers for user data collection as discussed below. These lines and numbers are drawn using the same feature of “SaveGrid” of the software developed in the first phase of this research, and described in chapter 3.

Recall that PET is the time from the end of encroachment of the left-turning vehicle to the beginning of encroachment of the conflicting through vehicle. The red lines represent the approximate wheel path of though vehicles. These paths are required to judge the end of encroachment point of left-turning vehicles. That is, the start of PET is when the rear bumper of the left-turn vehicle crosses the far (i.e. closest to edge of the road) red line of the conflicting through vehicle. The end of PET is the time at which the conflicting through vehicle enters the area of encroachment.

However, the turning path of left-turn vehicles may vary. It was determined to be inefficient to attempt to indicate where each left-turn vehicle left the area of encroachment. Therefore, a number of left-turning vehicles were observed and the general path followed by vehicles was derived based on user judgment. This gave the end of encroachment location for a majority of left-turning vehicles and was selected as the end of encroachment location for data collection. This location was used to draw the blue lines, to guide the data collector in the identification of the end of PET (i.e. time conflicting through vehicle enters conflict area). Though this selection was based on a judgment, it was informed by the observation of multiple vehicles. The blue numbers in Figure 4.1 act as a guide for noting the direction of movement of the involved vehicles. For example, Eastbound-Left is entered into the software as 62 (from: 6, to: 2), Eastbound-Through as 64 etc. A screenshot of the intersection of GA 10 with Henry Clower Blvd/Oak Rd is shown in Figure 4.2.

Figure 4.1: Intersection of GA 10 with Grayson Pkwy

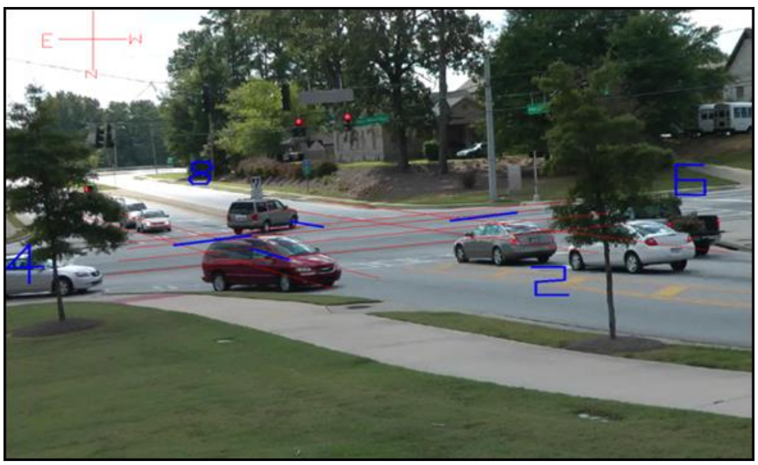


Figure 4.2: Intersection of GA 10 with Henry Clower Blvd/Oak Rd

4.3 RESULTS

This section presents the data and analysis for the two intersection pairs. First discussed is the off-peak data collected at the intersections of GA 10 with Grayson Pkwy and GA 10 with Henry Clower Blvd/Oak Rd., followed by a discussion of the peak hour data at these intersections. The analysis then proceeds to discuss the data collected at Buford

Hwy with Sugarloaf Pkwy and Roswell Rd with Wieuca Rd. Based on these results hypotheses that will guide the next phase of the research presented in chapter 5 will be presented.

4.3.1 Intersections of GA 10 with Grayson Pkwy, and GA 10 with Henry Clower Blvd/Oak Rd

As a surrogate to crash data, the PET data is expected to correlate in some manner with the crash data at some threshold PET value although a high correlation is not required to establish the diagnostic power of PET. It can be recalled from the literature review that in previous studies the selection of PET threshold has generally been chosen to be between 3 seconds and 5 seconds. A threshold is important as it is the boundary value below which PET is expected to exhibit its surrogate property. Therefore, according to previous studies, PET distributions at intersections having varied levels of safety should show differences below some threshold value at the lower end of the PET distribution.

The first PET surrogate to be considered is the ratio of the proportion of PETs under a given value, represented by the CDF. Since the total number of opposing left-turn conflicts differ between intersections, so will the number of PETs observed below a threshold. Therefore, it was decided to first investigate the proportion of PETs below a threshold. The absolute number of PETs (cumulative totals under a given threshold) is also considered as a potential surrogate.

4.3.1.1 PET Ratio as Potential Surrogate

As discussed, the first set of data for this intersection pair was collected on the 4th of October, 2010 during off-peak hours from 10 AM to Noon and again from 1 PM to 4 PM. The CDFs for the reduced PET data, as seen in Figure 4.3, show that there is no significant difference in measured PET at the two intersections, particularly at the lower end of the distribution, which is thought to be correlated to crashes. The crash data shows that the ratio of the number of all crashes which occurred at the intersection of Henry Clower Blvd/Oak Road and those at the intersection of Grayson Pkwy is 11:63, approximately 1:6. Thus, it would appear from this initial data that the proportion of low PETs is failing to reflect the crash data.

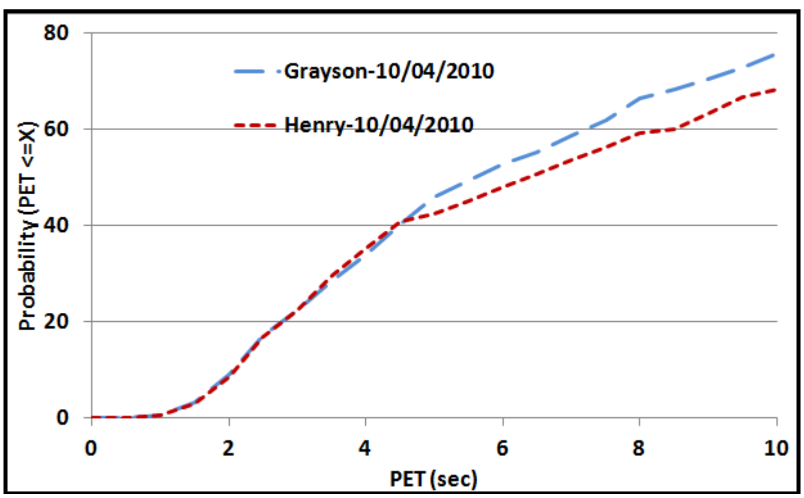


Figure 4.3: CDF plots of PET data collected on the 4th of October, 2010

To evaluate the consistency of this result the second set of non-peak data (collected April 7th, 2011) is shown in Figure 4.4. Though there is more divergence between the distributions of PET data collected in the second data set than the first, the distributions still significantly overlap each other at the lower tail of the distribution with PET values 3 seconds or less.

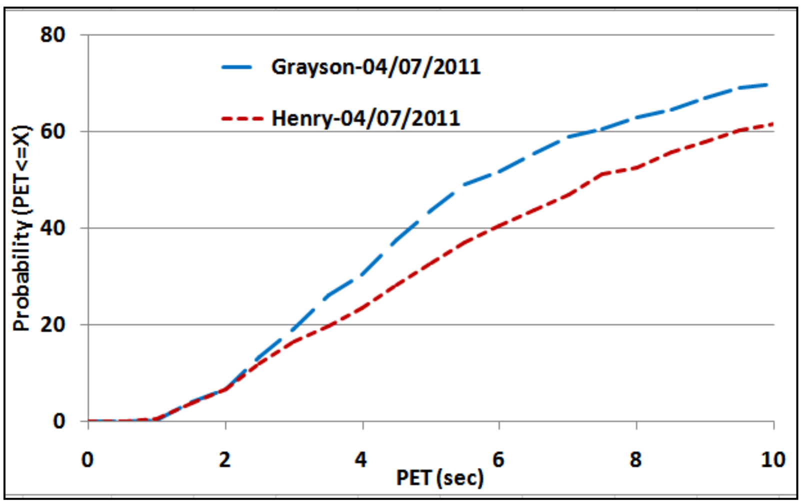


Figure 4.4: CDF plot of PET data collected on the 7th of April, 2011

It has already been mentioned in subsection 4.2.4 that based on data collected during non-peak hours, it was decided to collected peak-hour data for one day. Moreover, 30 percent of crashes occurred during peak hours. Figure 4.5 shows the comparison between peak and non-peak hours with respect to PET data and through traffic volume. It can be seen that there is considerable difference in PET data collected during peak and non-peak hours. Peak hour PET data mostly consist of low values, but the frequency of PET data at non-peak hour is higher. These observations led to the decision to collected PET data during a period having combination of peak and non-peak hours for intersections considered later. Figure 4.6 shows the peak-hour PET data collected from the two intersections on May 6th, 2011. This CDF data shows that Henry Clower Blvd/Oak Rd intersection has higher proportion of low PET values than those from Grayson Pkwy intersection. The expectation, at least from the definition of PET is that lower the PET value, higher the risk of crash. The expectation that follows is that higher the proportion of low PET values, higher the risk of crashes. By this logic, Henry Clower Blvd intersection (that had fewer crashes) was expected to have lesser proportion of low PET values than Grayson Pkwy intersection, but the data showed a distribution which is contrary to what might be expected from the crash data.

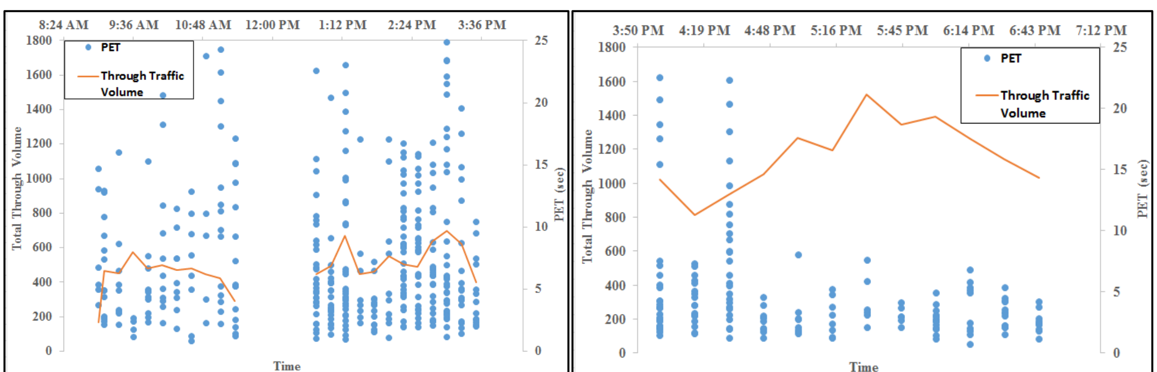


Figure 4.5: Peak (6th May, 2011) vs. Non-Peak (7th April, 2011) PET data and through traffic volume variation at the intersection of GA 10 with Grayson Pkwy

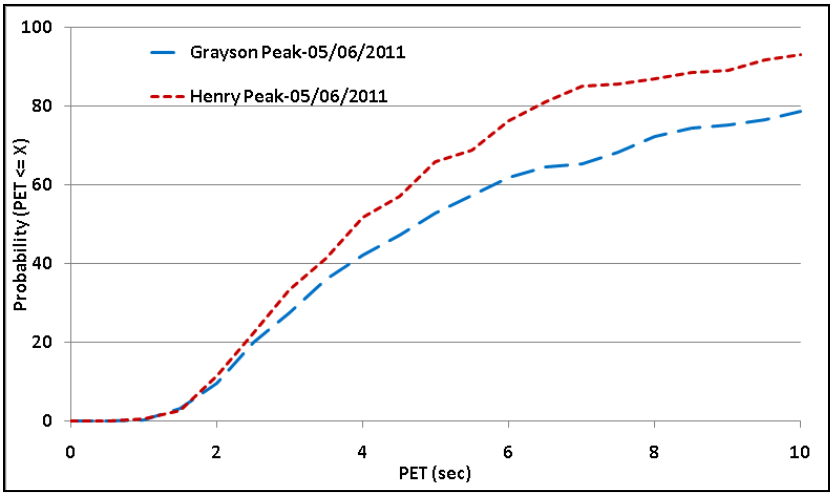


Figure 4.6: GA 10 with Grayson Pkwy, and GA 10 with Henry Clower Blvd/Oak Rd, CDF plots of PET data, Peak Period, 6th of May, 2011

4.3.1.2 Number of PETs as Potential Surrogate

Thus far, the CDF value of PET, which is the proportion of PETs below a threshold has been the primary PET surrogate considered. It was seen that this quantity measured at GA 10 with Grayson Pkwy, and GA 10 with Henry Clower Blvd/Oak Rd did not reflect the differences in crash data. In addition to the proportion of PETs, the quantity of absolute number of PETs below a threshold could be considered as another PET based surrogate measure. From the non-peak data, it can be seen that the number of PETs 3 seconds or less is similar for the Grayson Pkwy intersection and Henry Clower Blvd/Oak Rd intersection. If we extend the threshold to 6 seconds, the Grayson Pkwy intersection had higher number of PETs than Henry Clower Blvd/Oak Rd intersection but still not substantially different than that observed in the proportion of PETs below a threshold (Table 4.1). From the PM peak hour data, it can be seen that the number of PETs 3 seconds or less is higher for the Grayson Pkwy intersection than for Henry Clower

Blvd/Oak Rd intersection, which is different from that observed from the proportion of PETs (Table 4.2). This shows that a mere absolute cumulative frequency count reflects the same pattern as seen from crash data, though not in the same magnitude. Therefore, for these intersections it is seen that absolute cumulative frequency counts, though echoing the pattern found in the crash data, do not reflect the magnitude of this difference.

Table 4.4: Absolute frequency counts of the non-peak hour PET data

Grayson Pkwy - NonPeak				Henry Clower Blvd/oak Rd - NonPeak			
Bin	Frequency	Cum. Fre.	Cum. Prob.	Bin	Frequency	Cum. Fre.	Cum. Prob.
0	0	0	0.0	0	0	0	0.0
0.5	0	0	0.0	0.5	0	0	0.0
1	2	2	0.3	1	3	3	0.5
1.5	20	22	3.7	1.5	19	22	3.4
2	24	46	7.7	2	25	47	7.3
2.5	43	89	14.8	2.5	41	88	13.7
3	34	123	20.5	3	32	120	18.7
3.5	40	163	27.1	3.5	30	150	23.4
4	28	191	31.8	4	28	178	27.7
4.5	41	232	38.6	4.5	33	211	32.9
5	36	268	44.6	5	22	233	36.3
5.5	27	295	49.1	5.5	23	256	39.9
6	19	314	52.2	6	22	278	43.3
6.5	19	333	55.4	6.5	19	297	46.3
7	20	353	58.7	7	20	317	49.4
7.5	14	367	61.1	7.5	23	340	53.0
8	19	386	64.2	8	13	353	55.0
8.5	11	397	66.1	8.5	15	368	57.3
9	13	410	68.2	9	15	383	59.7
9.5	14	424	70.5	9.5	18	401	62.5
10	10	434	72.2	10	10	411	64.0
More	167	601	100.0	More	231	642	100.0

Table 4.5: Absolute frequency counts of peak hour PET data

Grayson Pkwy - Peak				Henry Clower Blvd/Oak Rd - Peak			
Bin	Frequency	Cum. Fre.	Cum. Prob.	Bin	Frequency	Cum. Fre.	Cum. Prob.
0	0	0	0.0	0	0	0	0.0
0.5	0	0	0.0	0.5	0	0	0.0
1	1	1	0.4	1	1	1	0.7
1.5	7	8	3.3	1.5	3	4	2.7
2	15	23	9.6	2	13	17	11.6
2.5	25	48	20.1	2.5	16	33	22.4
3	18	66	27.6	3	16	49	33.3
3.5	20	86	36.0	3.5	12	61	41.5
4	15	101	42.3	4	15	76	51.7
4.5	12	113	47.3	4.5	8	84	57.1
5	13	126	52.7	5	13	97	66.0
5.5	11	137	57.3	5.5	4	101	68.7
6	11	148	61.9	6	11	112	76.2
6.5	6	154	64.4	6.5	7	119	81.0
7	2	156	65.3	7	6	125	85.0
7.5	7	163	68.2	7.5	1	126	85.7
8	10	173	72.4	8	2	128	87.1
8.5	5	178	74.5	8.5	2	130	88.4
9	2	180	75.3	9	1	131	89.1
9.5	3	183	76.6	9.5	4	135	91.8
10	5	188	78.7	10	2	137	93.2
More	51	239	100.0	More	10	147	100.0

4.3.2 Intersections Wieuca Rd at Roswell Rd (in Fulton County) and Buford Hwy at Sugarloaf Pkwy (Gwinnett County)

This second intersection pair was chosen to have a more significant difference in the crash history, while still having similar geometry and traffic volumes. Screenshots are presented in Figure 4.7 (Wieuca Rd at Roswell Rd), and Figure 4.8 (Buford Hwy at Sugarloaf Pkwy).

Figure 4.7: Intersection of Roswell Rd and Wieuca Rd

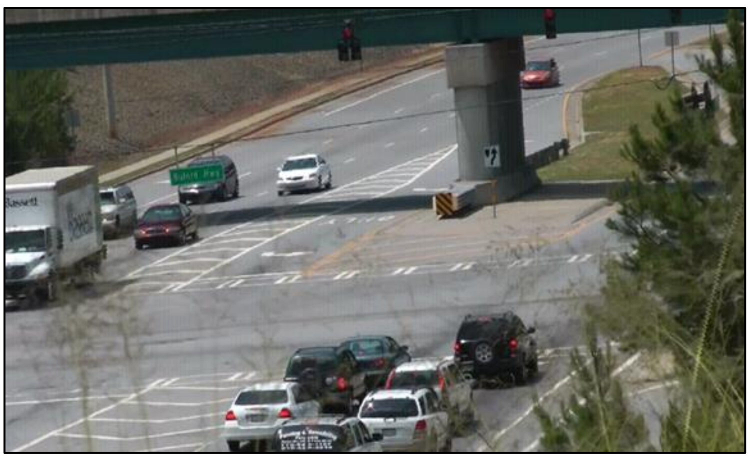


Figure 4.8: Intersection of Buford Hwy and Sugarloaf Pkwy

The intersection of Roswell Rd and Wieuca Rd had opposing left-turn crashes from the years 2006 through 2009. Comparing with other intersections and opposing left-turn crash frequencies, it can be said that this intersection may be considered to have a high potential for crashes with respect to the crash type being considered. The intersection of Buford Hwy and Sugarloaf Pkwy had only 6 left-turn opposing through crashes over the same 4 year period. The ratio of the total number of left-turn opposing through crashes at these intersections is approximately 1:13.

As seen in Table 4.3 video data at both these intersections was collected both during the peak and non-peak hours. Data at the intersection of Roswell Rd and Wieuca Rd was collected on 31st of May, 2011 from 7 AM to 12 noon while data at the intersection of Buford Hwy and Sugarloaf Pkwy was collected on 3rd of June, 2011 from 2 PM to 7 PM. There were some logistic reasons due to which the data could not be collected during the AM peak period at the intersection of Buford Hwy and Sugarloaf Pkwy. The video data was reduced to obtain PET data (Figure 4.9). An investigation of the PET data collected at these two intersections shows that approximately 25% of PET values recorded at the intersection of Roswell Rd and Wieuca Rd are less than or equal to 3 seconds whereas only 4% of PET data collected at the intersection of Buford Hwy and Sugarloaf Pkwy are less than or equal to 3 seconds, a very pronounced difference. Data also shows that for a PET value of 1 second or less, the cumulative probability value for the intersection of Buford Hwy and Sugarloaf Pkwy is 0.202 while that for the intersection of Roswell Rd and Wieuca Rd is 2.72 which is approximately 14 times greater. As shown in Figure 4.10, the PET value increases, this factor decreases to a factor of approximately 5 and is never greater than the factor of 14 at a PET value of 1 second. It is also observed that there are no PET values below 0.5 seconds at both the intersections.

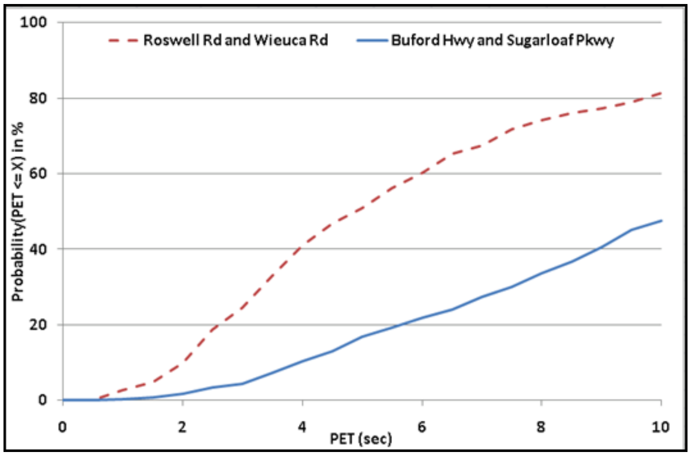


Figure 4.9: CDF plots of PET data collected at the intersections of Roswell Rd with Wieuca Rd, and Buford Hwy with Sugarloaf Pkwy

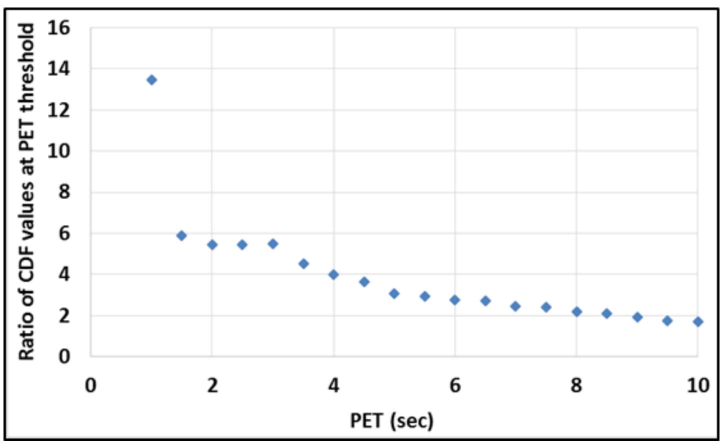


Figure 4.10: Ratio of CDF values at PET thresholds between data of the intersections of Roswell Rd with Wieuca Rd, and Buford Hwy with Sugarloaf Pkwy

Similarly, as presented in the first set of intersections, the absolute number of PETs observed below a threshold value could be considered as a representative of crash propensity, as shown in Table 4.6.

Table 4.6: Absolute frequency counts of PET data at the intersections of Roswell Rd with Wieuca Rd, and Buford Hwy with Sugarloaf Pkwy

Roswell Rd. and Wieuca Rd				Buford Hwy and Sugarloaf Pkwy			
Bin	Frequency	Cum. Fre.	Cum. Prob.	Bin	Frequency	Cum. Fre.	Cum. Prob.
0	0	0	0.0	0	0	0	0.0
0.5	0	0	0.0	0.5	0	0	0.0
1	8	8	2.7	1	1	1	0.2
1.5	6	14	4.8	1.5	3	4	0.8
2	15	29	9.9	2	5	9	1.8
2.5	26	55	18.7	2.5	8	17	3.4
3	17	72	24.5	3	5	22	4.4
3.5	25	97	33.0	3.5	14	36	7.3
4	24	121	41.2	4	15	51	10.3
4.5	17	138	46.9	4.5	13	64	12.9
5	12	150	51.0	5	19	83	16.8
5.5	15	165	56.1	5.5	12	95	19.2
6	12	177	60.2	6	13	108	21.8
6.5	15	192	65.3	6.5	11	119	24.0
7	6	198	67.3	7	17	136	27.5
7.5	13	211	71.8	7.5	13	149	30.1
8	7	218	74.1	8	18	167	33.7
8.5	6	224	76.2	8.5	15	182	36.8
9	3	227	77.2	9	19	201	40.6
9.5	5	232	78.9	9.5	22	223	45.1
10	7	239	81.3	10	12	235	47.5
More	55	294	100.0	More	260	495	100.0

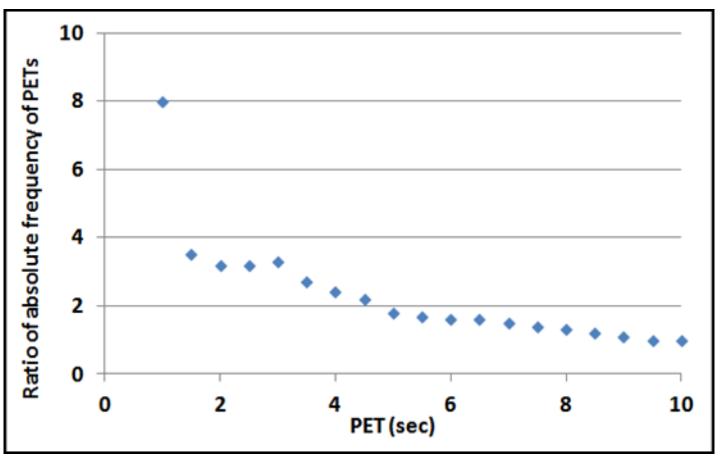


Figure 4.11: Ratio of absolute frequency counts between PET data of the intersections of Roswell Rd with Wieuca Rd, and Buford Hwy with Sugarloaf Pkwy

The intersection of Roswell Rd and Wieuca Rd had 72 interactions which had a PET value of 3 seconds or less whereas the intersection of Buford Hwy and Sugarloaf Pkwy had 22 interactions in this PET range. Therefore, the absolute frequency of PETs below a threshold of 3 seconds shows a factor of 3.27. Figure 4.11 shows this factor for each threshold value.

For a PET value of 1 second or less, the intersection of Roswell Rd and Wieuca Rd has 8 times more observations than those from the intersection of Buford Hwy and Sugarloaf Pkwy. As the threshold value increases, this ratio factor decreases and for a PET value of 3 seconds or less, the factor is 3.3.

4.4 SUMMARY

This phase of the research delved deeper into the evaluation of PET as a surrogate measure for safety with respect to the interactions between left-turn vehicles and opposing through vehicles. Two pairs of intersections, having varied levels of safety but having similar operating conditions were selected for this phase of research. The first pair of intersections (Grayson Pkwy and Henry Clower Blvd/Oak Rd, both with GA 10) had a ratio of 3:1 in terms of crashes, while the second pair ((Roswell Rd with Wieuca Rd and Buford Hwy with Sugarloaf Pkwy) had a ratio of 13:1 crashes. While this phase of research did not provide any concrete evidence of the effectiveness of PET, it led to the development of a set of questions that need to be answered in the third research phase where PET is collected at a number of additional intersections.

Question 1

The first pair of intersections (Grayson Pkwy and Henry Clower Blvd/Oak Rd, both with GA 10) has a ratio of approximately 3:1 for the left-turn opposing through crashes which occurred from the year 2000 to 2009 while the second pair of intersections (Roswell Rd with Wieuca Rd and Buford Hwy with Sugarloaf Pkwy) has a ratio of approximately 13:1 for the same type of crash over the same time period. The analysis of the PET data collected at these intersections showed that the Grayson Pkwy and Henry Clower Blvd/Oak Rd intersections did not show a significant difference in the PET data distribution. The intersections of the Roswell Rd and Wieuca Rd, and Buford Hwy and Sugarloaf Pkwy on the other hand show significant differences in the PET data collected. Therefore, it can be hypothesized that PET, as a surrogate measure of safety, may act as an effective surrogate for comparing safety of two intersections having high differences in crashes while it may not capture the difference between intersections which are moderately different in safety or that significantly more data is needed to capture such differences.

Question 2

This analysis has also led to a consideration that PET alone cannot act as a sufficient surrogate. The following relationship between PET and crashes will therefore be considered in the next chapter.

Critical event

$$P(\text{crash}) = P(\text{PET} <= \text{threshold}) \times P(\text{crash/critical event})) \tag{4.1}$$

Analysis of PET data would provide the first part of this equation which gives the probability of a PET value being less than a threshold. In other words, the PET data gives the probability of having a critical event. But the second part of the equation provides for the likelihood of a critical event leading to a crash. That is, PET alone does not provide the probability of a crash which means it alone cannot act as a surrogate for crashes. PET combined with other parameter(s) (e.g. traffic volume, sight distance, grade etc.) gives the probability of a critical event leading to a crash. Therefore this combination of PET and parameters may act as a surrogate. In fact, it may be possible that a quantitative relationship between crashes and PET data in terms of predicting the number of crashes or correlation with crashes cannot be established and that PET may only be used as a qualitative measure for determining relative crash propensities of intersections or as a diagnostic tool.

Question 3

The third question that needs to be addressed is the representation of PET for a surrogate measure of safety. For example, further research needs to be performed to determine if the absolute number of PETs below a threshold correlate with crashes or if the proportion of PETs below a threshold is a better surrogate. It is also possible a different form of PET would be reasonable. It can also be seen that there may be a requirement for a larger sample size to make concrete conclusions. As a part of this phase of the research, data from two pairs of intersections were collected. Though these intersections had different levels of crash frequency and traffic volumes, data from a wider range of intersections having different crash frequencies and traffic volumes would give a larger sample size allowing for conclusions about the effectiveness of PET as a surrogate measure for safety to be drawn with a higher confidence.

The next phase of the research discussed in chapters 5 and 6 seeks to address these questions.

CHAPTER 5: ANALYSIS AND MODEL DEVELOPMENT

This chapter begins with a descriptive analysis of the available crash and PET data and their underlying distributions. This analysis will inform potential modeling techniques considered later in this chapter to model crash frequency in terms of PET and other intersection characteristics. As discussed in Chapter 3, some previous studies have hinted at the potential use of surrogates as a diagnostic measure for potential crash risk rather than as a predictive measure for crash frequency. Similarly, Chapter 4 presented additional hypotheses on the effectiveness and application of PET as a surrogate measure for safety that resulted in additional data collection. Building on these efforts, this chapter examines the PET data collected at these study intersections. This analysis examines both the statistical properties of the PET data and the extent to which the PET data can be used either as a non-crash-based safety modeling technique to predict crash probabilities or as a diagnostic factor for identifying high or low risk intersections.

5.1 DESCRIPTIVE STATISTICS OF CRASH DATA

As a first step in this analysis, the distribution of observed crashes associated with an opposing left-turn conflict, our study conflict, at intersections in the state of Georgia for the years 2006 through 2009 was determined. Due to logistic reasons, it was decided to limit the study intersections to Atlanta area, and therefore corresponding distribution for Atlanta area was also determined. This distribution of crash frequencies provides a context for the analysis performed in the later phases of this research. Accident records for the years 2006 through 2009 were provided by GDOT and processed to evaluate candidate intersections for the PET study. Relevant data were extracted based on the RCLINKs from crash records and the *Critical Analysis Reporting Environment* (CARE) software (*23*), as discussed in chapter 4.

All the intersections selected for the study are located in Atlanta, GA metropolitan statistical area (MSA) and were selected to have varied levels of observed safety (defined below) to ensure that any models developed to differentiate between intersections having different propensities for opposing left turn crashes could be tested across a broad range of conditions. A histogram of observed opposing left-turn crashes, based on an initial database search using RCLINKs, for the selected intersections during years 2006-2009 is presented in Figure 5.1. The distribution of opposing left-turn crash frequencies across intersections in Georgia shown in Figure 5.1 is heavily skewed. The vast majority of Georgia intersections (≈ 89%) have five or fewer crashes over the four year period 2006-2009 and only a very small fraction (≈ 0.1%) of intersections have greater than 40 crashes in the four years with the remaining intersections (≈ 11%) having a frequency of crashes between 6 through 40. It can be noted that intersections having not even one crash in the 4-year period are not listed in the CARE database. So, the lowest interval of crashes in this distribution is 1-5.

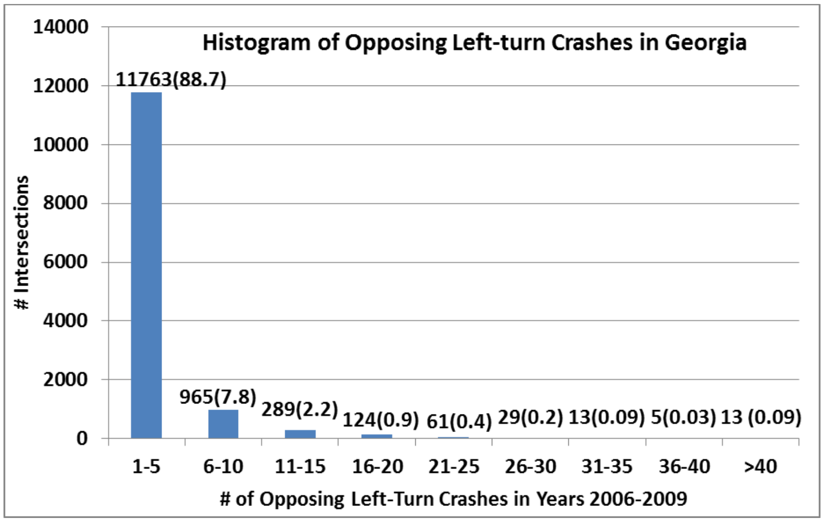


Figure 5.1: Histogram of Opposing Left-Turn Crashes in Georgia

There are three criteria that were taken into consideration for selecting the study intersections:

- The overall goal of this phase of the research is to evaluate the effectiveness of PET as a surrogate measure to differentiate intersections having varied levels of observed safety (as reflected by crash numbers). It was concluded from phase 2 of the research (chapter 4) that intersections having a wider range of crash frequencies need to be considered in this phase of research to evaluate the effectiveness of PET. Since the distribution of crashes is so skewed, sufficient number of intersections from the extreme right tail needed to be considered to obtain the required range of crash frequencies. Hence it is important to have a mix of intersections having varied levels of crashes, assuming PET and crashes have the same mechanism, and that needs to be evaluated.
- The "major road AADT" was another criterion for selection of intersections as the CARE crash database included only major road AADT among its data fields. Only

intersections having a major road AADT of greater than 20,000 were considered for this study to control for the effect of variation in AADT (exposure). As stated previously it is expected that one of the major applications of PET will be to determine the effectiveness of any safety treatment or countermeasures applied at an intersection without having to wait the typical 3 year period for collecting crash data, limiting before versus after AADT variation. Thus, to evaluate the effectiveness of PET as a treatment MOE, it is advantageous to select intersections which have similar operating conditions but different crash frequencies. In addition, the majority of intersections having a major road AADT less than 20,000 tended to be unsignalized intersections or biased towards having very few opposing left-turn crashes (attributed to very low conflicting volumes). Hence only those intersections having major road AADT of greater than 20,000 were considered. Once the intersections were selected, STARS (State Traffic and Report Statistics) 2011 data from GDOT was used to obtain updated major and minor road AADTs.

- As mentioned already, for logistic reasons, it was decided to selected intersections only from the Atlanta 6-county (Fulton, Gwinnett, Cobb, Dekalb, and Clayton, Rockdale) metro area.

Given these criteria, the distribution of crash frequencies was again determined. Figure 5.2 presents this distribution. The number of intersections having between 1 and 5 crashes has fallen from 11763 to just 975. On the other hand, there are still 11 intersections having greater than 40 crashes (one of the 13 intersections in Georgia having greater than 40 opposing left-turn crashes is in Hall county, and the other has a major AADT less than 20,000).

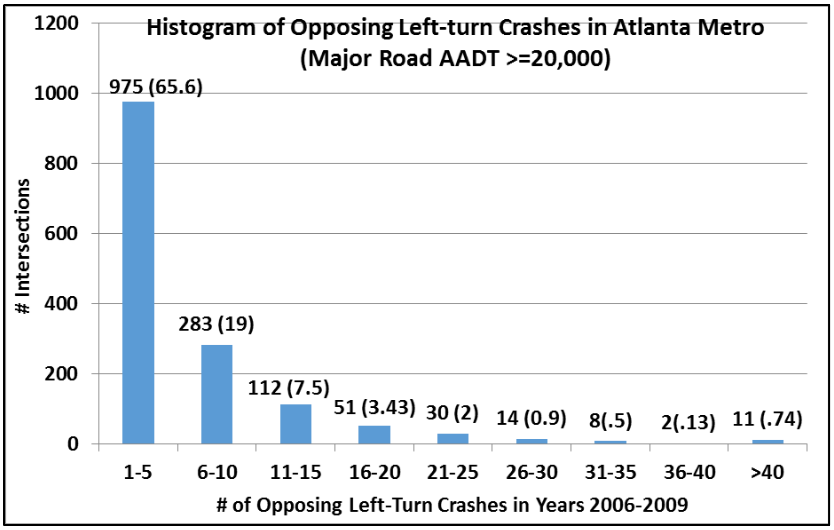


Figure 5.2: Histogram of Opposing Left-Turn Crashes in Atlanta Metro area

In order to evaluate the diagnostic power of PET (its power to classify an intersection as having a high or low potential for opposing left-turn crashes), study intersections should be selected such that they could be classified into high or low crash categories based on crash numbers. There is no theoretical basis behind considering any particular crash frequencies as thresholds for differentiating between high, medium, and low crash frequency categories. Given that the distribution in Figure 5.2 is still heavily skewed, and since the frequency of the interval 1-5 crashes is very high compared to the other intervals, this interval was segregated into a separate category. Since this category represents intersections that have less than or equal to 5 crashes in 4-years, it is named "low" category. Intersections having greater than 35 crashes in 4-years (9 crashes per year) make up 1% in the histogram. Hence intersections having greater than or equal to 10 crashes per year were considered to be in the "high" category, making 40 opposing left-turn crashes as the threshold for "high" category for 4-year crash counts. The intersections having 4-year opposing left-turn crash counts between 6 and 39 are therefore considered to be belonging to "medium" category. It is possible that there is no medium category, and that the category 6-39 is actually "high" category, and >40 crashes is a "very high" category. But, just to follow the common nomenclature, it was decided to name the category 6-39 as "medium" crash intersections.

Limitations of time and resources restricted the study to 18 intersections and these were selected having a significance variance in crash frequency (based on the initial RCLINK based search of the crash data), from as many as 75 (highest identified) to as few as one over the four year period, and a mixture of hazard levels with respect to left-turn opposing crashes. In order to have a mixture of intersections with varying levels of safety with respect to left-turn opposing crashes, 6 intersections were selected from each category mentioned above.

Although arbitrary, the exact classification scheme (later corrected in absolute crash values) should not significantly affect the results of subsequent analyses. Analyses that involve evaluating the "classification" ability or diagnostic power of PET can be affected due to the selection bias (due to classifying intersections based on crash frequency). Even here, the majority of impact potentially would be on the "medium" category, which could in reality have been belonging to "high" or "low" category. At the end of this chapter, section 5.6 presents an analysis that explores the impact of classification bias on establishing the diagnostic power of PET parameters. This bias however would not have any effect for analyses involving rank correlations or parametric modeling where absolute crash numbers are used.

Once the intersections were selected, PET data was collected in the same methodology as applied in phase 2 of the research. As already mentioned in section 4.2.2, the first potential problem with crash data was identified during this phase of research and section 5.2 below describes how the collected PET helped to identify this problem.

5.2 POTENTIAL ROLE OF PET IN RECOGNIZING ERRORS IN CRASH DATA

PET data were collected at the study intersections using the methodology described earlier. The intersection of Lawrenceville Hwy and N. Druid Hills Rd (subsequently referred to as intersection #1 in Figures and Tables), based on the initial RCLINK based raw crash data, was shown to have 75 crashes listed between the years 2006 and 2009 and stood as the intersection having the highest number of crashes between left-turn vehicles and opposing through vehicles in the state of Georgia, as seen in the CARE software. PET data were collected at this intersection and the cumulative numbers of PETs observed in each half-second threshold category at and below three seconds in comparison to other intersections is presented in Table 5.1. At this initial stage of the analysis the PET threshold for recognizing a serious conflict was not known and an initial three second threshold was considered as it was SSAM's default threshold value.

It is generally assumed that the lower the PET value, the greater the severity of conflict and proximity to a crash. From the Table it can be seen that though the total number of PETs (439) observed at Lawrenceville Hwy and N. Druid Hills Rd is almost the highest, only 74 PETs out of a total of 439 PETs (16.8%) recorded fell at or below 3 seconds while other intersections that had much lower crash frequencies had higher percentage of PETs at or below 3 seconds. Moreover, only 2 of the PETs were at or below 1 second (a very low PET value and hence a conflict of high potential seriousness). Regarding the absolute number of PETs the intersection of Lawrenceville Hwy and N. Druid Hills Rd also does not have the greatest values at any of the PETs at or below 3 seconds. Therefore, overall trends in the PET data collected at Lawrenceville Hwy and N. Druid Hills Rd did not provide qualitative agreement with the crash data. This potentially undermines either the usefulness of PET data or potentially indicates a problem with the underlying crash data.

Table 5.1: PET data at high crash intersections, based on RCLink Crash Database Search

Intersection	Crashes	Total # PETs	# of PETs				
			<=3 sec	<= 2.5 sec	<=2 sec	<=1.5 sec	<=1 sec
N Druid Hills Rd and Lawrenceville Hwy	75	439	74	47	26	12	2
Roswell Rd and W Wieuca Rd	57	434	72	55	29	14	8
Lawrenceville Hwy and Lawrenceville Suwanee Rd	52	241	53	34	19	8	2
GA 138 and Sigman Rd	46	183	40	24	13	9	8
GA 20 and Willow Lane	44	444	121	98	67	38	10
Grayson Hwy and Scenic Hwy	40	312	83	64	38	23	6

Based on these results, it was decided to verify the crash data at this and all other study intersections by direct comparison using police reports as a quality assurance measure. These police reports revealed that 42 crashes out of the 75 crashes attributed to Lawrenceville Hwy and N. Druid Hills Rd actually occurred at another intersection – N. Druid Hills Rd and Lavista Rd (intersection # 2). Crash data also shows that N. Druid Hills Rd and Lavista Rd had 27 crashes (investigation of police reports revealed that 3 of these were also incorrectly attributed). With the addition of 42 crashes attributed incorrectly to Lawrenceville Hwy and N. Druid Hills Rd, the total number of crashes at N. Druid Hills Rd and Lavista Rd should have been stated as 66 (based on crashes originally identified using an RCLINK based search), resulting in this intersection with a higher number of crashes than those intersections under study. Though this intersection was not considered initially, after this crash data verification, PET data were then collected at this intersection and added to the analysis.

A comparison of the PET distributions obtained based on the data collected at these two intersections (as shown in Table 5.2) and the corrected initial intersection (although still based on the RCLINK based search) crash counts demonstrated a broad agreement in intersection safety categorization. While this is a single example of no statistical significance, it has shown the potential ability of PET to indicate crash probability, potentially creating an independent means to screen for errors in crash data which if successful could be a significant application of the PET method. Most transportation funding agencies rely on the crash data to rank intersections and to fund projects. The above analysis shows that PET may be able to serve as a tool to guide decision makers and increase their confidence in recognizing intersections that require safety treatments and funding.

Table 5.2: Comparison of PET data collected at intersection # 1 and intersection # 2

Intersection	Cummulative # PETs									
	<=10 s	<=8 s	<=6 s	<=4 s	<=3.5 s	<=3 s	<=2.5 s	<=2 s	<=1.5 s	<=1 s
Intersection # 1	310	265	206	120	97	74	47	27	12	2
Intersection # 2	181	165	131	87	67	61	45	28	16	9

However, as mentioned in chapter IV, additional issues with the crash data were also identified. It has already been shown that some of the crashes were wrongly attributed to intersections due to errors in assigning the RCLink. However, at a later time in the research, "sanitized" databases for the study years were provide by GDOT which had intersection road names (major and cross road) instead of the RCLink and milepost combination. Since CARE software identified some of the crashes correctly, this new data set has been used to find out the additional crashes that were missing from the RCLink-based search. Using this new data, crash records for the study intersections (searching on the road names) were again filtered and verified for any data that were not obtained from CARE software. Police reports for these crashes were again referred to for filtering out the records corresponding to opposing left-turn crashes, and also to identify any other errors/inaccuracies. It was found that through the use of the road names in the sanitized databases additional records for the study intersections were identified.

5.2.1 Final Crash Data

All the intersections considered in this study are located in Atlanta, GA metropolitan statistical area (MSA) and are shown in the map (Figure 5.3).

Figure 5.3: Study area (Source: www.maps.google.com)

All the data used for the analysis in the succeeding sections are the corrected data based a road name search of the sanitized database and validated through comparison with police reports. Table 5.3 shows the frequency of crashes (after verification) between left-turning vehicles and opposing through vehicles for the years 2006 through 2009 at the 18 intersections selected for this study. No statement could be made on the state rankings of these intersections as it has already been seen that the raw crash data has many errors and the corrections have been made only for the study intersections. It is important to highlight that the updated crash data has increased the number of crashes identified at almost all intersections compared to the initial RCLINK based search. However, the above discussion regarding the use of PET data to potentially identify errors in crash data continues to hold, although with limitations. The updated crash data did not change the relatively ranking (by absolute number of crashes) between the intersections as identified previously. Thus, the PET data is potential useful in identifying this potential ranking error but not generally useful in identifying an error in the absolute number of incidents. Since the crash frequencies for the study intersections have generally increased, it was decided to shift the ranges for “high”, “medium”, and “low” categorizations. Hereafter, intersections having 0-10 crashes are considered to be belonging to “low” crash category, 11-50 crashes belong to “medium” crash category, and greater than 50 belong to “high” category. Even with the new categorization, each category still has 6 intersections.

Table 5.3: Study intersections and crashes

Intersection	Crashes based on road name search (2006-9)
N Druid Hills Rd and Lavista Rd	118
GA 138 and Sigman Rd	90
Roswell Rd and W Wieuca Rd	78
Lawrenceville Hwy and Lawrenceville Suwanee Rd	73
GA 20 and Willow Lane	64
Grayson Hwy and Scenic Hwy	53
N Druid Hills Rd and Lawrenceville Hwy	48
GA 10 and Grayson Pkwy	29
Ponce De Leon Ave and Moreland Ave	27
Scott Blvd and Clairemont Ave	23
Memorial Dr and Covington hwy	15
Glenwood Dr and Columbia Dr	15
GA 10 and Oak Rd	9
Sugarloaf Pkwy and Buford Hwy	6
Cobb Pkwy and Gresham Rd	5
MLK Jr Dr and Brownlee Rd	2
Whitlock Ave and Lindley Ave	2
North Ave and Techwood Dr	2

5.3 PET DATA ANALYSIS

PET data were collected at the 18 study intersections using the methodology described in section 4.3. The descriptive statistics for Post Encroachment Time (PET) measurements at the study intersections are presented for all data and PET observations of less than three seconds in Tables 5.4 and 5.5 respectively. These data were collected over a period of 5 hours from 2 PM to 7 PM on a weekday at each intersection. The total numbers of PETs recorded varied from 87 to 495 with mean observed PET values ranging between 5.56 sec and 10.3 sec.

Table 5.4: Descriptive statistics of the PET data collected at the 18 study intersections

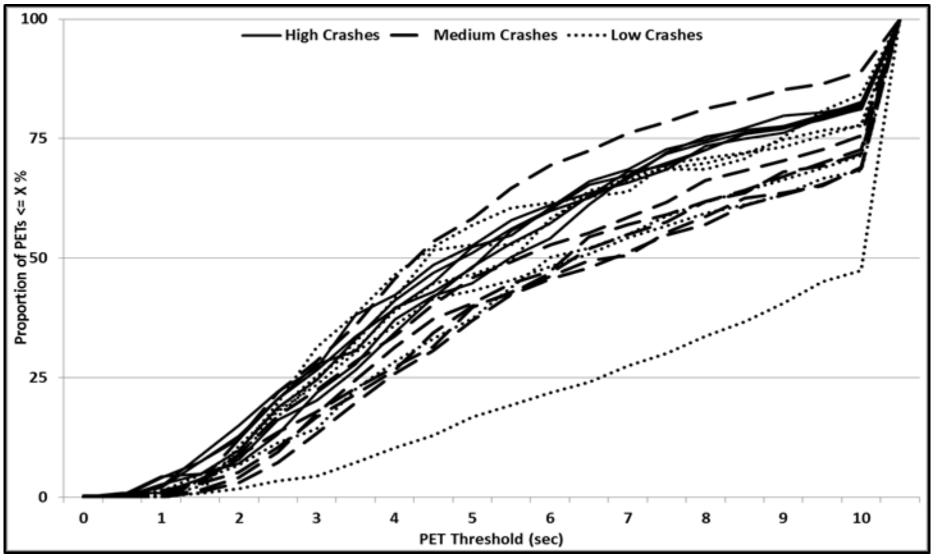
Intersection	Crashes	Count	Mean	Median	Mode	Std. Dev.	Kurtosis	Skewness
N Druid Hills Rd and Lavista Rd	118	219	8.25	6.57	3.1	5.36	4.44	1.94
GA 138 and Sigman Rd	90	183	6.7	5.47	3.53	5.86	29.5	4.15
Roswell Rd and W Wieuca Rd	78	434	6.12	4.88	3.5	4.26	0.87	1.15
Lawrenceville Hwy and Lawrenceville Suwanee Rd	73	241	6.54	5.2	3.37	5.05	15.02	2.83
GA 20 and Willow Lane	64	444	6.35	4.73	1.96	5.27	4.97	1.94
Grayson Hwy and Scenic Hwy	53	312	6.46	4.83	3.33	5.4	4.79	2.01
N Druid Hills Rd and Lawrenceville Hwy	48	439	7.97	6.24	4.37	5.88	3.99	1.71
GA 10 and Grayson Pkwy	29	464	7.15	5.22	6.7	6.57	40.7	4.41
Ponce De Leon Ave and Moreland Ave	27	331	5.56	4.32	1.88	5.66	87.6	7.75
Scott Blvd and Clairemont Ave	23	368	9.24	6.5	4	7.82	1.6	1.44
Memorial Dr and Covington Hwy	15	359	9.48	6.75	3.62	10.36	71.62	6.52
Glenwood Dr and Columbia Dr	15	338	7.53	6.24	2.25	5.15	3.36	1.42
GA 10 and Oak Rd	9	340	9.06	4.83	2.02	16	21.11	4.54
Sugarloaf Pkwy and Buford Hwy	6	495	10.3	10.2	15	5.01	1.8	0.73
Cobb Pkwy and Gresham Rd	5	420	8.69	6.39	3.03	7.48	8.28	2.4
MLK Jr Dr and Brownlee Rd	2	87	6.95	4.19	2.06	6.09	2.06	1.61
Whitlock Ave and LindleyAve	2	104	7.81	5.5	1.76	8.06	16.4	3.33
North Ave and Techwood Dr	2	89	6.41	4.43	2.3	4.64	1.47	1.34

Table 5.5: Distribution of PET data at 3 second lower thresholds, and crash counts

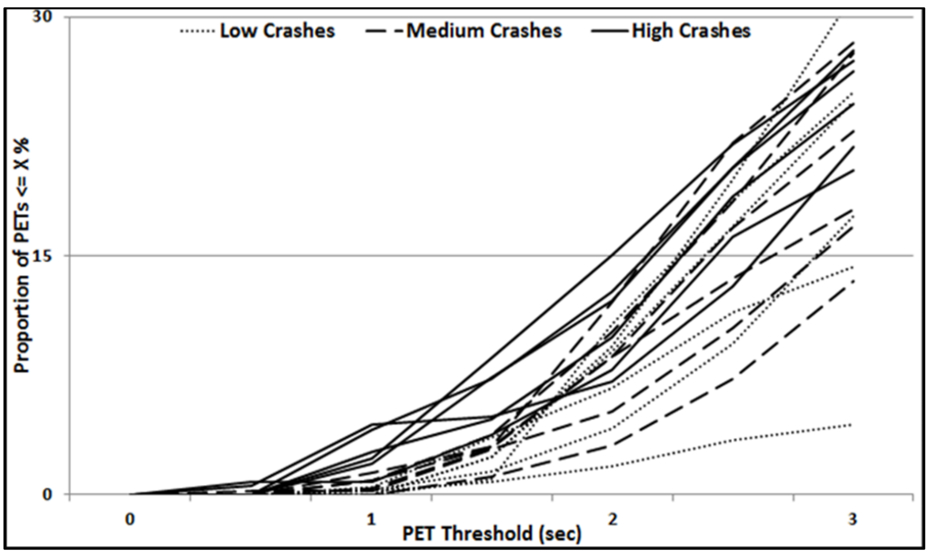
Intersection	# PETs						Crashes
	Total	<=3 sec	<= 2.5 sec	<=2 sec	<=1.5 sec	<=1 sec	
N Druid Hills Rd and Lavista Rd	219	61	45	28	16	9	118
GA 138 and Sigman Rd	183	40	24	13	9	8	90
Roswell Rd and W Wieuca Rd	434	72	55	29	14	8	78
Lawrenceville Hwy and Lawrenceville Suwanee Rd	241	53	34	19	8	2	73
GA 20 at Willow Lane	444	121	98	67	38	10	64
Grayson Hwy and Scenic Hwy	312	83	64	38	23	6	53
N Druid Hills Rd and Lawrenceville Hwy	439	74	47	26	12	2	48
GA 10 and Grayson Pkwy	464	106	73	44	15	2	29
Ponce De Leon Ave and Moreland Ave	331	94	73	40	10	3	27
Scott Blvd and Clairemont Ave	368	66	50	32	11	5	23
Memorial Dr and Covington Hwy	359	48	26	11	4	0	15
Glenwood Dr and Columbia Dr	338	63	41	16	6	2	15
GA 10 and Oak Rd	464	95	63	35	13	3	9
Sugarloaf Pkwy and Buford Hwy	495	22	17	9	4	1	6
Cobb Pkwy and Gresham Rd	420	60	48	28	15	1	5
MLK Jr Dr and Brownlee Rd	87	27	17	8	2	0	2
Whitlock Ave and LindleyAve	104	26	19	11	1	0	2
North Ave at Techwood Dr	89	19	13	6	2	0	2

Distribution of any measure may often be depicted by Cumulative Distribution Function (CDF) plots. CDFs were plotted (Figure 5.4 (a)) to compare the distribution of PETs corresponding to the observed safety levels of the intersections. The solid black lines represent intersections having a high number of crashes (greater than or equal to 50); the broken lines represent intersections having a medium number of crashes (greater than 10 but less than 50); and the dotted lines represent intersections having a low number of crashes (less than or equal to 10). An examination of this Figure leads to an initial broad conclusion that the CDF plots of PET data collected at various intersections overlap, demonstrating little correspondence to crashes. However, it may be seen that a potential pattern at lower levels of PET thresholds may exist. Moreover, previous studies have suggested that a PET value as high as 3 seconds as a potential threshold. Figure 5.4 (b) expands these plots up to the 3 second values.

Figure 5.4 (b) shows that at a threshold of 3 seconds, the three intersection categories do not follow a clear pattern. Evidence of any pattern would suggest an ability of PET to classify an intersection(s) into any of the three categories of hazard. However, this plot suggests that PET likely does not have such a capability at a threshold of 3 seconds, a threshold suggested by many previous studies. However, at lower PET thresholds there appears to be an increasing potential for PET to indicate higher crash locations. The next section will explore this possibility.



(a)



(b)

Figure 5.4: CDF plots of PET data collected at study intersections (a) Complete dataset, and (b) An expanded view of CDF plots truncated at a PET of 3 seconds.

5.4 UTILITY OF PET DATA IN CRASH PREDICTION

From the definition of PET, it can be understood that a PET value of 0 would imply the occurrence of a crash. It follows that determining the distribution of PET values and its associated parameters can lead to a predicted crash propensity (i.e. the probability of PET being 0). If this is possible, PET data can be directly used to predict crashes (or rankings) at an intersection(s) independent of the crash history at the location (except for validation). In other words, to develop models to predict crash frequencies without direct access to crash data.

To understand how this is possible, we will resort to some statistical arguments (many of which can be found in standard statistical texts like Coles (2001), and Ang and Tang (2006)). Let us suppose that the distribution PET data is given by F(t) where t is the PET value. This means that computing the value of F(0) would represent the probability of a crash. This is, of course, the rationale for using PET as a surrogate for determining crash frequency. Hence, it is important to understand or determine the distribution that the PET data follow at any specific location.

Any statistical distribution can be defined by its probability density function (PDF) or, equivalently, its cumulative distribution function (CDF) and their corresponding parameters. For our purposes, a cumulative distribution function is particularly useful. As a CDF gives the probability that a random variable X is less than a given value x it is natural way of describing thresholds (i.e. a value that we would be interested if an observation were either greater (or less) than). Standard "textbook" CDF are generally based on a theoretical analytical framework. On the other hand, an empirical distribution function (EDF) is obtained from the data rather than a theoretical formulation. The word empirical means "obtained from experiment" and as such uses data obtained from experiment or field to estimate the distribution of a random variable. This distribution is a step function which increases by a value of 1/n for each of the n data points. Let $(x_1,\ldots\ldots,x_n)$ be real random variables with the common CDF F(X). The empirical distribution function is defined as (Ang and Tang, 2006):

$$F_n(x) = (\text{number of values in the data sample} <= x)/n = (1/n)* \sum_{i=1}^{n} I\{x_i <= x\} \tag{5.1}$$

where I{} is the indicator function.

This analysis uses the observed EDF to estimate the underlying distribution of PET data rather than using a purely theoretical framework.

For each of the study intersections, crash numbers for four years and PET data from 5 hours of data collection for one day is available. Consider one of the study intersections as an example. The intersection of N Druid Hills Rd and Lavista Rd recorded 118 crashes in four years (2006 through 2009). For this intersection, a part of distribution of PET data collected is shown in Table 5.6.

Table 5.6: PET distribution at the intersection of N Druid Hills Rd and Lavista Rd

Number of PETs observed <=				
3 sec	2.5 sec	2 sec	1.5 sec	1 sec
61	45	28	16	9

These crashes form the "tail" of the PET distribution as a crash is implied by a PET value of 0. While the distribution of complete PET dataset is unknown, the probability of a crash can be considered as a conditional probability based on a threshold. To see how this works, let us consider a threshold of 3 seconds. Thus the probability of the occurrence of a PET value of 0 (i.e. a crash) can been expressed conditionally based the occurrence of PET of 3 seconds or less. That is:

$$P(PET <= 0) = \frac{PET <= 0}{PET <= 3sec} * P(PET <= 3sec) \tag{5.2}$$

Let us take for the moment that P(PET<=3sec) is a constant or follows some distribution and focus on the conditional probability. Let us continue with the example of the intersection of N Druid Hills and Lavista Rd. The distribution of PET values at this intersection is already shown the Table 5.3. Since the number of PET values <=0 (crashes) is available only as a number for 4 years, we have to compute the number of potential occurrences of PET values less than or equal to 3 seconds in 4 years.

PET values observed less than or equal to 3 sec in 5 hrs of data collection = **61**

PET values less than or equal to 3 sec potentially observed in a day = 61*3 = **183**

(assumes conflicts over 15 hours, discounting for very low volume hours.)

PET values less than or equal to 3 sec potentially observed in 4 years = 183*365*4 = **237980**

Therefore, P(PET<=0/PET<=3 sec) = 118/237980 = **0.00049583**

Likewise, computing this conditional probability for PET values given PET<= 3 is also straightforward.

P(PET<=1sec/PET<=3sec) = 9/61 = 0.1475409

P(PET<=1.5sec/PET<=3sec) = 16/61 = 0.26229508

P(PET<=2sec/PET<=3sec) = 28/61 = 0.459016

P(PET<=2.5sec/PET<=3sec) = 45/61 = 0.73770491

P(PET<=3sec) = 1

Similarly, for the same intersection, conditional probabilities can be computed using different threshold values of 2.5s, 2s, and 1.5s. This process of computing conditional CDF values can be repeated for other study intersections and the resulting values can be seen in Appendix A .

Generalized Extreme Value Distribution

The Generalized Extreme Value Distribution is a family of three continuous probability distributions (Gumbel, Frechet, and Weibull) that have their base in extreme value theory. Extreme value theory is a concept in statistics that is used to model occurrences whose probability is extremely deviant from the median probability of distribution (Coles (2001)). This theory is generally applied to model rare events such as earthquakes, 100-year floods, etc. The expression for CDF of GEV distribution is

$$F(x; \mu, \sigma, \xi) = \exp\{-[1+\xi((x-\mu)/\sigma)]^{(-1/\xi)}\} \tag{5.3}$$

where,

$\mu \in R$ = location parameter

$\sigma > 0$ = scale parameter

$\xi \in R$ = shape parameter

for $1+\xi((x-\mu)/\sigma > 0$

The shape parameter ξ governs the tail behavior of the distribution. The sub-families defined by $\xi = 0$, $\xi > 0$ and $\xi < 0$ correspond, respectively, to the Gumbel, Fréchet and Weibull families. Table 5.7 shows the parameters of the GEV distribution fitted to the data at each PET threshold level 4 seconds and below at 0.5 seconds interval.

Table 5.7: PET distribution at the intersection of N Druid Hills Rd and Lavista Rd

Threshold	Location	Scale	Shape
4	1.93	0.91	-0.33
3.5	1.86	0.83	-0.44
3	1.79	0.78	-0.54
2.5	1.53	0.66	-0.53
2	1.19	0.50	-0.46
1.5	0.88	0.32	-0.30
1	0.74	0.21	-0.65

The location parameter shifts to lower PET values, which is expected because the data points on which the distribution is being fit, is limited to the threshold value. The scale of the distribution also seems to be decreasing, and decreasing more rapidly as the threshold value decreases. This means the CDF plot becomes flatter. The shape parameter does not follow any pattern as the threshold varies. The plots in Figure 5.5 show the comparisons between the observed and computed CDF values based on the fit distribution for data below each threshold.

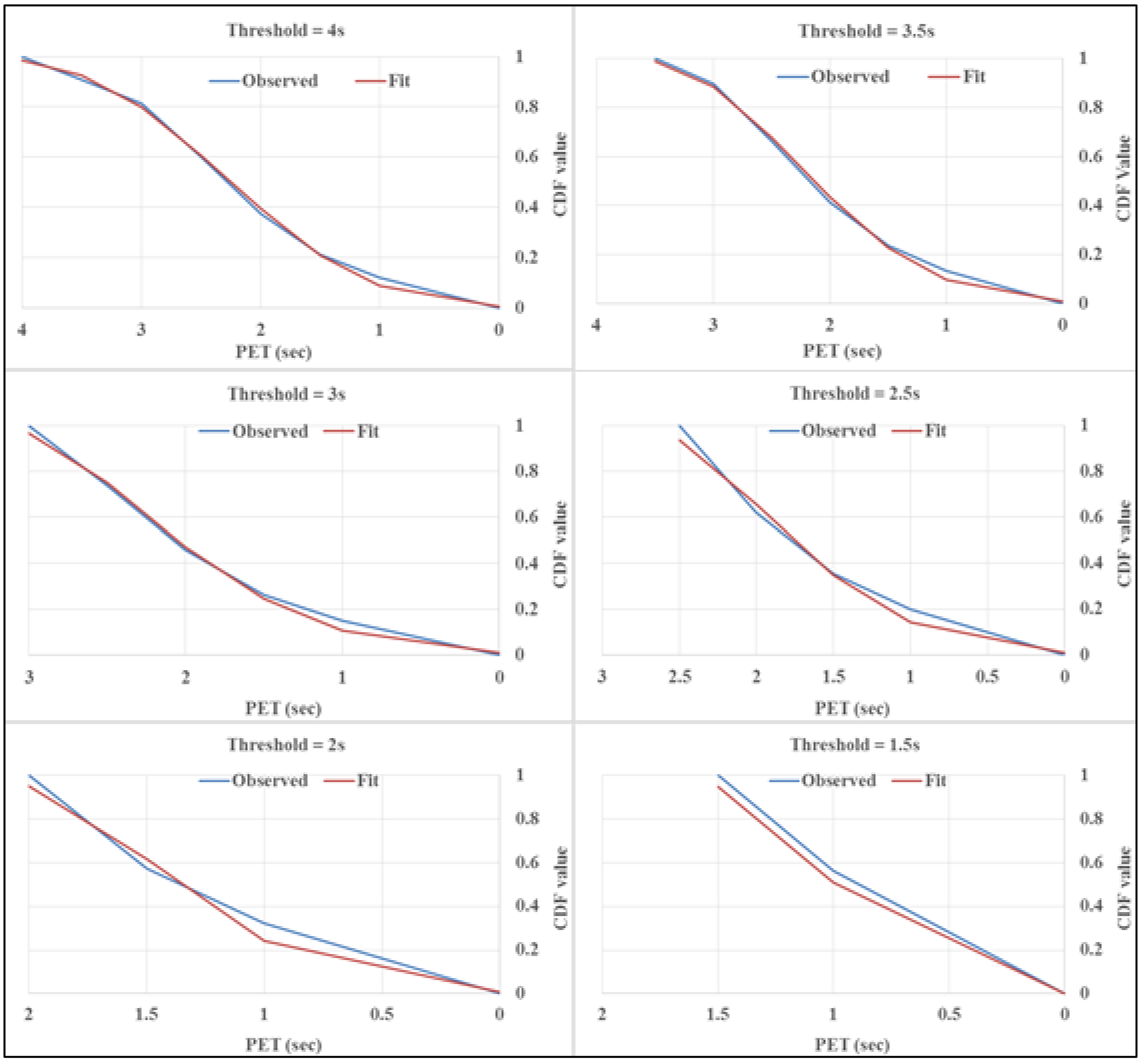


Figure 5.5: GEV fit for the PET data collected at the intersection of N Druid Hills Rd and Lavista Rd

All the plots show that though there is good agreement between fit and observed CDF values. At the plotted scale it appears the observed and fit crash probabilities (CDF value at 0 second) match vary closely. However, CDF value (probability) is a multiplicative measure and hence looking at just the absolute values gives a distorted picture of the scenario. It is the ratio of the predicted CDF value to the observed CDF value that which is of interest rather than the absolute difference. A closer look at the data values shows actual differences at a PET value of 0. Table 5.8 shows these comparisons.

Table 5.8: Comparison of the observed and fit crash probabilities (PET=0)

Threshold	Observed	Fitted	Ratio
4	0.00019	0.00647	34.05
3.5	0.00021	0.00873	41.57
3	0.00026	0.01121	43.12
2.5	0.00032	0.01066	33.31
2	0.00051	0.00680	13.33
1.5	0.00090	0.00060	0.67

The Table shows that the ratio of the predicted to observed crash probabilities is approximately 40 at higher threshold values. Since ratios are of interest, the difference can better be understood by taking logarithm of the CDF values. The plots in Figure 5.6 show these values.

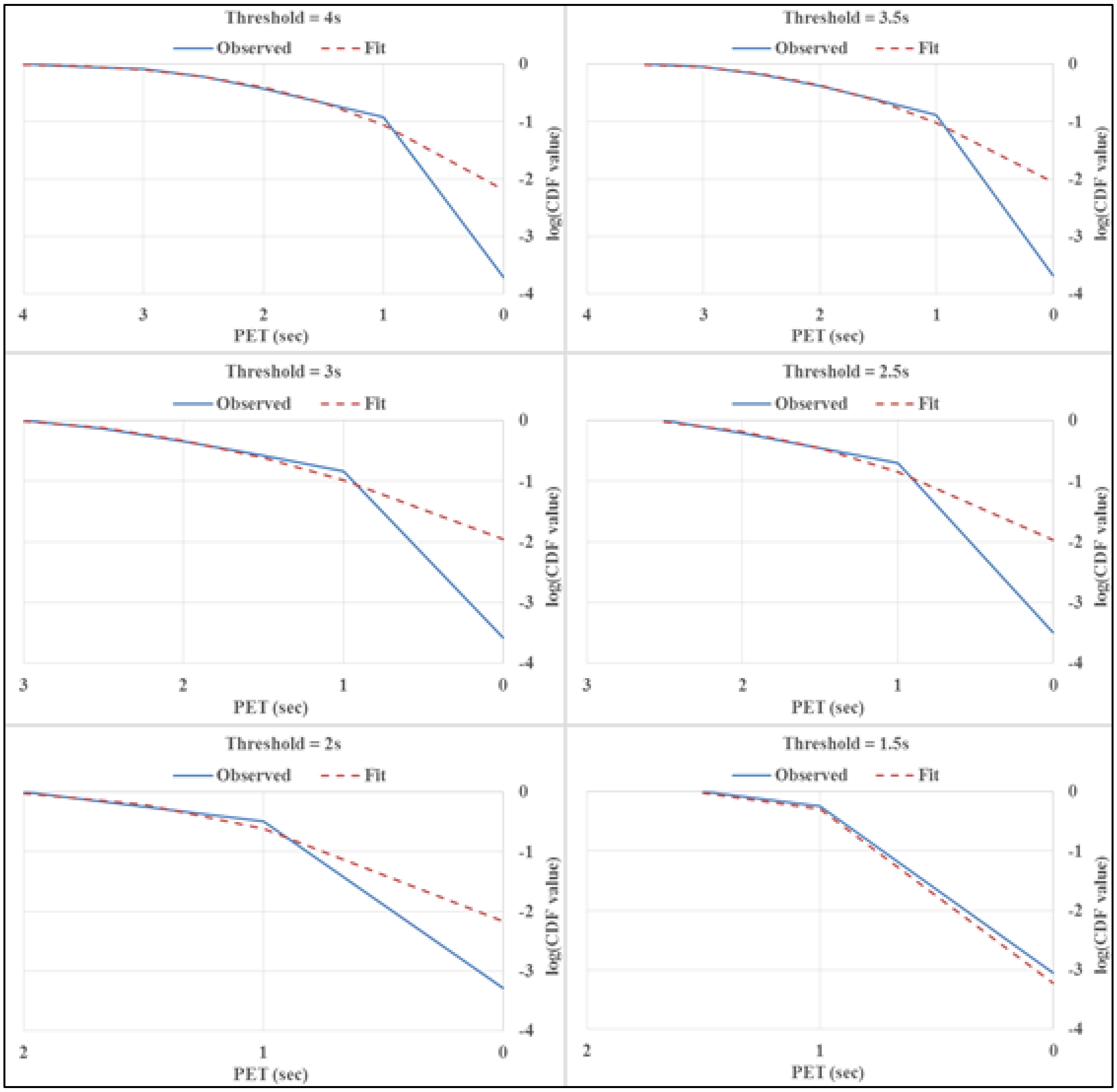


Figure 5.6: Plots of fit and observed CDF values on logarithmic scale for GEV fit of PET data collected at the intersection of N Druid Hills Rd and Lavista Rd

It can be seen that in most of the plots, predicted crash probabilities are more than the observed probabilities or in other words, crashes are over estimated. However, there is a significant decrease in this difference when only the data below 1.5 seconds is considered to fit the distribution. There are two observations that can be made from the data shown here. First of all, divergence between expected CDF values from GEV fit and observed

CDF values are considerable for 1 second and 0 sec (crash) when any threshold value 2 seconds and higher is considered to fit the GEV distribution. Second, the difference between observed crash probability and predicted crash probability decreases significantly as threshold value decreases to 1.5. One possible explanation is that this means that the process of PET occurrence of higher PETs (especially 2 seconds and more) is different from that of 1.5 seconds and below (likely due to driver intervention to reduce the possibility of a crash). However, this argument cannot be extended to infer that the GEV fit at 1.5 threshold is the same curve that crash occurrence follows. This is because for a GEV fit at 1.5 second threshold, there are only two data points (1.5s, and 1.0s). The only inference that can be made is that for GEV fits at thresholds greater than 1.5s, there is divergence between the observed, and fit probabilities at 1 second and 0 second (crash), which means that thresholds greater than 2.0 s are not fit to predict crashes, probably due to driver intervention at lower PET values. This trend is seen in the data collected at other study intersections too. Corresponding plots for the other intersections are shown in the Appendix B.

At a threshold of 1.5s, on the other hand, there are only two data points (1.5 second and 1 second) which limits concluding about the data's ability to fit PET=0. Moreover due to unavailability of sufficient PET data below 1 second, it is not possible to determine the distribution at a threshold of PET=1 second and below. For some of the study intersections, there are no data points below 1 second, and in some cases, very few below 1.5 seconds, which inhibits fitting a GEV or any other distribution at such a threshold. In such cases, the same conclusions cannot be extended, until more data is collected in future research efforts and similar trends are observed. In summary, using a threshold above 1.5 seconds does not follow the correct distribution to predict crashes, and at or below a threshold of 1.5. seconds, the correct conclusion can only be drawn if sufficient data and data points to fit the distribution were existing.

This analysis tries to understand the PET data distribution in detail and exposes that to predict crashes using just PET data (and not even crash history), obtaining more PET data at 1 second threshold and below is very important. This is an important finding because it shows that a PET threshold such as 3 seconds that is generally used in such studies may not be sufficient to predict crashes or evaluate safety. PET data considered at higher thresholds tend to overestimate crashes and it is clear that the process of crash occurrence is very different to the occurrence of conflicts that have PET values of 2 and above. It also shows that some other type of modeling approach that involves crash numbers needs to be used to predict crashes. These modeling approaches are discussed in chapter 6.

5.5 NON-PARAMETRIC DATA ANALYSIS

Rank correlation measures the degree of similarity between variables in terms of similarity of the rankings. It is a very important non-parametric statistical measure of relationship between two variables that is often used when the actual values of these variables are not as important as the relative ordering or ranking of their values (Carterette (2009)). However, it should be understood that for this study, the actual values (crash frequencies) are very important and rank analysis only supplements the knowledge gained from using parametric modeling approaches. Rank analysis is performed here to find out the strength of PET in acting as a diagnostic measure vis-a-vis its performance as a predictor. This correlation is measured in terms of a correlation coefficient. The rank correlation coefficients often used are:

- Spearman's ρ
- Kendall's τ
- Goodman and Kruskal's γ

The value of the rank correlation coefficient ranges between [-1, 1]. The higher the value of this coefficient, the better the agreement between the variables in terms of rankings. Spearman's ρ and Kendall's τ are particular cases of a generalized correlation coefficient that is calculated as (Kvam and Vidakovic (2007))

$$\rho = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}. \tag{5.4}$$

This generalized relationship is especially useful when there are ties in the rankings in which case the formulas for Spearman's ρ and Kendall's τ cannot be applied. In this study, there are numerous occasions where there were ties in the ranks. Hence the generalized form of rank correlation coefficient has been used. Moreover, whenever there is a tie, identical values are assigned a rank equal to the average of their positions in the ascending order of the values.

Two types of rank correlation analyses are performed. The first considers all data and computes the rank correlation coefficient to determine the overall correlation between total crashes and the considered parameter. To understand the methodology to perform this analysis, an example case is shown in Table 5.9. This example considers the relationship between major road AADT and corresponding crash frequencies by computing rank correlation coefficient. As it can be seen, there are four values of 1 and they are assigned a rank that is the average of position values of all 1s i.e. (13+14+15+16)/4 = 14.5.

Table 5.9: Computation of rank correlation coefficient between number of crashes and major road AADT

Intersection	Total AADT	Rank (x)	(x-avg(x))^2	Crash Frequency	Rank (y)	(y-avg(y))^2	(x-avg(x))* (y-avg(y))
N Druid Hills Rd and Lavista Rd	48805	6	12.25	118	1	72.25	29.75
GA 138 and Sigman Rd	44715	7	6.25	90	2	56.25	18.75
Roswell Rd and W Wieuca Rd	38200	10	0.25	78	3	42.25	-3.25
Lawrenceville Hwy and Lawrenceville Suwanee Rd	43490	8	2.25	73	4	30.25	8.25
GA 20 and Willow Lane	31380	16	42.25	64	5	20.25	-29.25
Grayson Hwy and Scenic Hwy	54547	2	56.25	53	6	12.25	26.25
N Druid Hills Rd and Lawrenceville Hwy	49695	5	20.25	48	7	6.25	11.25
GA 10 and Grayson Pkwy	52253	4	30.25	29	8	2.25	8.25
Ponce De Leon Ave and Moreland Ave	54540	3	42.25	27	9	0.25	3.25
Scott Blvd and Clairemont Ave	59650	1	72.25	23	10	0.25	-4.25
Glenwood Dr and Columbia Dr	39545	9	0.25	15	11.5	4	-1
Memorial Dr and Covington Hwy	35385	14	20.25	15	11.5	4	9
GA 10 and Oak Rd	36130	12	6.25	9	13	12.25	8.75
Sugarloaf Pkwy and Buford Hwy	36710	11	2.25	6	14	20.25	6.75
Cobb Pkwy and Gresham Rd	35820	13	12.25	5	15	30.25	19.25
Whitlock Ave and LindleyAve	33230	15	30.25	2	17	56.25	41.25
North Ave and Techwood Dr	24360	17	56.25	2	17	56.25	56.25
MLK Jr Dr and Brownlee Rd	23710	18	72.25	2	17	56.25	63.75
			Sum			**Sum**	**Corr. Coeff.**
			484.5			**482**	**0.56**

However, the overall correlation may be guided by a strong relationship in one section of intersections (high crash or low crash intersections). One parameter might be good at filtering out intersections having very low crashes but may show weak performance in differentiating between intersections having medium and high crash propensity. Another parameter might have completely opposite strength in which case it can be used to find if an intersection falls in the high crash category. Therefore, the second type of analysis tries to determine the performance of parameters with respect to their effectiveness in differentiating between intersections having high and low number of crashes.

The task here is a categorical classification to test if a parameter can classify intersections into the three categories. Given that the number of samples belonging to each category is small, instead of conducting a three-way classification test, two-way classification test can also be conducted by merging any two categories and testing it against the third. For example, the medium and high crash categories can be merged and a classification test can be conducted to verify if the parameter can filter out low crash intersections vs. medium-high. Similar test can be conducted to test high vs. medium-low. If a parameter can do both, it indicates an ability to classify intersections into all three categories. Given the number of samples in each category, a Fischer's exact test is used (Kvam and Vidakovic (2007)) instead of a chi-square test. Table 5.10 (a) shows the set up for the experiment.

Table 5.10: (a) Fischer's test for the parameter PET_1s (High-Medium vs. Low) (b): Resulting contingency table

Crashes	Category	PET_1s	Category
64	H	10	H
118	H	9	H
78	H	8	H
90	H	8	H
53	H	6	H
23	H	5	H
27	H	3	H
9	L	3	H
73	H	2	H
48	H	2	H
29	H	2	H
5	L	2	H
15	H	1	L
6	L	1	L
15	H	0	L
2	L	0	L
2	L	0	L
2	L	0	L

(a)

	Crash	PET_1s
H	10	12
L	2	0

(b)

Since the aim here is to test if the parameter can classify an intersection appropriately, first sort the intersections in a descending order of the parameter (example in Table 5.10 (a) is PET_1s). The test is to verify if PET_1s can successfully classify those intersections that belong to the high-medium category (and thereby the remaining in low category). Since high and medium are being combined, the first 12 intersections for PET_1s, all belong to H category (means no intersection belongs to L category). Corresponding categories based on crash numbers are also shown. The contingency Table can be derived as shown in Table 5.10 (b).

The null hypothesis of a Fischer's exact test is that the two groups (crash and PET_1s) are statistically not different. The one-tailed p-value of the test based on the Table 5.10

(b) equals 0.2391 which means there is no evidence to reject the null hypothesis. This means that PET_1s can be used to filter out high-medium intersections (in other terms low crash intersections).

Repeating the experiment to compare high vs. medium-low, gives a set-up as shown in 5.11 (a) and corresponding contingency Table in 5.11 (b)

Table 5.11: (a) Fischer's test for the parameter PET_1s (High vs. Low-Medium) (b) Resulting contingency table

Crashes	Category	PET_1s	Category
64	H	10	H
118	H	9	H
78	H	8	H
90	H	8	H
53	H	6	H
23	L	5	H
27	L	3	L
9	L	3	L
73	H	2	L
48	L	2	L
29	L	2	L
5	L	2	L
15	L	1	L
6	L	1	L
15	L	0	L
2	L	0	L
2	L	0	L
2	L	0	L

(a)

	Crash	PET_1s
H	1	0
L	11	12

(b)

This test gives a p-value of 0.5, again no evidence to reject the null hypothesis that PET_1s can be used to classify intersections belonging to low-medium category (which means high category as well). Since, the tests show that PET_1s could be used to classify high-medium vs. low as well as high vs. medium-low, it can be inferred that PET_1s parameter can classify intersections into all three categories.

This test only tells if a parameter has the ability to classify. However, it does not tell the comparative nature of the classification (i.e., if a parameter classifies high from medium-low better than the other way around, or comparison of performance between parameters). A straight-forward measure of agreement between the ranks of crashes and a parameter can be computing the average of the absolute difference between the ranks (AADR) of crashes and the parameter. This measure considers the corresponding difference in rank values in the set of intersections and computes an average difference. The computation of this measure is shown in Table 5.12. From the definition of AADR, it can be seen that lower the AADR, better is the agreement between the ranks of crash and the parameter.

Table 5.12: Computation of AADR

Intersection	Total AADT	Rank (x)	Crash Frequenc	Rank (y)	Abs. Diff. in	AADR
Scott Blvd and Clairemont Ave	59650	1	23	10	9	9.00
Grayson Hwy and Scenic Hwy	54547	2	53	6	4	6.50
Ponce De Leon Ave and Moreland Ave	54540	3	27	9	6	6.33
GA 10 and Grayson Pkwy	52253	4	29	8	4	5.75
N Druid Hills Rd and Lawrenceville Hwy	49695	5	48	7	2	5.00
N Druid Hills Rd and Lavista Rd	48805	6	118	1	5	5
GA 138 and Sigman Rd	44715	7	90	2	5	5
Lawrenceville Hwy and Lawrenceville Suwanee Rd	43490	8	73	4	4	4.88
Glenwood Dr and Columbia Dr	39545	9	15	11.5	2.5	4.61
Roswell Rd and W Wieuca Rd	38200	10	78	3	7	4.85
Sugarloaf Pkwy and Buford Hwy	36710	11	6	14	3	4.68
GA 10 and Oak Rd	36130	12	9	13	1	4.38
Cobb Pkwy and Gresham Rd	35820	13	5	15	2	4.19
Memorial Dr and Covington Hwy	35385	14	15	11.5	2.5	4.07
Whitlock Ave and LindleyAve	33230	15	2	17	2	3.93
GA 20 and Willow Lane	31380	16	64	5	11	4.38
North Ave and Techwood Dr	24360	17	2	17	0	4.12
MLK Jr Dr and Brownlee Rd	23710	18	2	17	1	3.94

5.5.1 Volume Based Analysis

Accident research has frequently relied on the concept of exposure, and different variations and definitions for it have been proposed (Chapman (1973), Hauer (2005), Elvik (2009)). Two of the most common measures of exposure used are vehicle miles travelled, and number of vehicles traveling through a location. The current research deals with the number of vehicles entering the intersection as a preliminary measure of exposure. Exposure can also be divided into two categories: summary measures, and elementary units (Elvik (2009)). Traffic volume is one of the most frequently used factors of safety representing exposure, and intuition suggests that the greater the exposure, the greater the likelihood for a crash. It is also understood that traffic volume only explains a portion of the variability in crash counts at various locations. The majority of crash prediction models suggested in previous studies (McDonald (1953), David and Norman (1975), Hakkert and Mahalel (1978), Hauer (2004)), including those in the Highway Safety Manual (2011), have used Annual Average Daily Traffic (AADT) counts in models to predict total crashes at an intersection. However, AADT is an aggregate measure of exposure and provides for an average opportunity for total crashes at any location. Elementary units of exposure on the other hand refer to events that create a more direct opportunity for a crash. From the videos recorded for collecting PET data, traffic counts were also collected for through vehicles and left-turning vehicles for all four approaches to an intersection. Left-turning vehicles were also divided into two categories - turning left on a protected phase and turning left on a permitted phase. In this effort VNBTh, VEBTh, VSBTh, VWBTh are the through vehicle volumes corresponding to northbound, eastbound, southbound, and westbound directions respectively, and VNBL, VEBL, VSBL, VWBL are the left-turning volumes on the permitted phase at each intersection corresponding to northbound, eastbound, southbound, and westbound directions, respectively. Since the conflict under consideration in this study is between left-turn and opposing through vehicles, conflicting traffic volumes instead of AADTs are a more pertinent measure of exposure. These numbers provide for a direct opportunity for crash as compared to AADTs. Previous studies (Pickering et al. (1986), Songchitruksa and Tarko (2004), Elvik (2009)) suggest that a product of through volume and opposing left-turning volume provides a measure of conflicting volumes. To allow for the same unit as "volume", a square root of this product is considered to be "conflicting volume". Therefore, the following equation would compute the total conflicting volume at an intersection.

$$Vc = \sqrt{(VNBTh*VSBL + VEBTh*VWBL + VSBTh*VNBL + VWBTh*VEBL)} \qquad (5.5)$$

Where

Vc = conflicting volume

This section analyses the agreement between crashes and volume measures (AADT, conflicting volume etc.) in terms of rank order. The computation of RCC between major road AADT and major road crashes is already shown in Table 5.9. Similar calculations were performed for the relationship of total crashes with total AADT, minor road crashes with minor AADT, and pairs of parameters corresponding to conflicting volumes. Please note that the crash numbers here and in all the analysis ahead are opposing left-turn crashes only. The results of these calculations are as follows:

Table 5.13: Rank correlation coefficients between number of opposing left-turn crashes and traffic volume measures

Variables	Rank Correlation Coefficient
Major road crashes and major AADT	0.15
Minor road crashes and minor AADT	0.52
Total crashes and total AADT	0.56
Major road crashes and major conflicting volume	0.76
Minor road crashes and minor conflicting volume	0.45
Total crashes and total conflicting volume	0.72

The highest rank correlation was found between major road crashes and major road conflicting volume with a coefficient of 0.76. Total crashes also show a high rank correlation with total conflicting volume with a coefficient of 0.72. However, with respect to AADT, minor road crashes had the best rank correlation with the corresponding AADT values resulting in a coefficient of 0.56 while low correlations were found for total crashes and major road crashes with the corresponding AADTs. This implies that if one wants to rank intersections in terms of propensity of crashes (instead of using crash numbers directly), conflicting volume may be a stronger parameter than AADT.

An important aspect to look at is the difference in the performance of these parameters in differentiating between high, medium, and low crash category. As explained before, first Fischer's exact test was done to test if AADT parameters have the ability to classify intersections into the three categories.

- Major AADT: The p-value for the test for High-Medium vs. Low category classification is 0.046 which means there is sufficient evidence to reject the null hypothesis. The p-value for Medium-Low vs. High is also 0.046. So, Major AADT is not a sufficiently good measure to classify the study intersections.
- Minor AADT: The p-value for the test for High-Medium vs. Low category classification is 0.5 which means there is not sufficient evidence to reject the null hypothesis. The p-value for Medium-Low vs. High is also 0.046. So, Major AADT has the ability to filter out low crash intersections but is not a sufficiently good measure to filter high intersections.
- Total AADT: The p-value for the test for High-Medium vs. Low category classification is 0.23 which means there is not sufficient evidence to reject the null hypothesis. The p-value for Medium-Low vs. High is also 0.046. So, Total AADT also has the ability to filter out low crash intersections but is not a sufficiently good measure to filter high intersections.

Since only Minor AADT, and Total AADT showed evidence of ability to classify low crash intersections, we can compute the AADR value to determine the quality of classification. The AADR values considering 6 lowest ranked intersections in terms of Minor AADT, and Total AADT are 3.16, and 3.08 respectively.

For all other parameters considered, the results are summarized in Tables that follow. Each Table consists of three columns. The first column is the parameter being studied for its ability to identify high crash and/or low crash intersections. The second column shows the p-value from Fischer's exact test for the comparison between High-Medium vs. Low category while the third column shows the p-value from Fischer's exact test for the comparison between Medium-Low vs. High category. There is value in comparing the AADR values only if a parameter exhibits the ability to classify intersections. Therefore AADR values would be discussed only based on the results of these tests.

The second parameter being considered in this analysis is conflicting volume. First of all, the values in Table 5.14 show that conflicting volume measures have similar ability to that of AADT in classifying intersections. All the three parameters show inability to classify high crash intersections while there is evidence to show that both major conflicting volume and total conflicting volume could be used to identify low crash intersections. This analysis shows that conflicting volume is a better measure for identifying low crash intersections than high crash intersections which agrees with that found for AADT. The corresponding AADR values for major conflicting volume, and total conflicting volume are 3.08 and 2.85. Though these are slightly lower than those found for AADT, it still remains to be seen if these AADR values are the lowest possible among all parameters.

Table 5.14: Results of Fisher's exact test using conflicting volume parameters

Parameter	p-value from Fischer's test	
	High-Medium vs. Low	High vs. Medium-Low
Minor Conflicting Volume	0.046	0.046
Major Conflicting Volume	0.23	0.046
Total Conflicting Volume	0.23	0.046

5.5.2 PET Based Analysis

The correlation between the rankings of CDF values at different PET thresholds and corresponding crash frequencies is shown in Table 5.15. This shows that CDF1s has the highest rank correlation with crash frequencies with a coefficient of 0.82 followed by CDF1.5s with a coefficient of 0.76. Though these values are higher those found for conflicting volume parameters, it is not a statistically significant difference.

Table 5.15: Rank correlation coefficients for the proportion of PETs below a threshold

Parameter	Rank Correlation Coefficient
CDF3s and # crashes	0.16
CDF2.5s and # crashes	0.31
CDF2s and # crashes	0.30
CDF1.5s and # crashes	0.76
CDF1s and # crashes	0.82

The correlation between the ranks of number of PETs observed below different thresholds (for ex. PET_1s) and corresponding crash frequencies is shown in Table 5.16.

Absolute number of PETs below 1 second shows the highest RCC with crash frequencies with a coefficient of 0.80 which is similar to the corresponding value observed for the parameter CDF1s and this is the highest RCC observed until this point in the analysis. However, it can be seen that RCC value for CDF1.5s is 0.75 which is much higher than the corresponding RCC value for number of PETs below 1.5 seconds. In case of insufficient data of PET = 1 second, CDF1.5s is a comparatively good measure to rank intersections.

Table 5.16: Rank correlation coefficients for the number of PETs below a threshold

Parameter	Rank Correlation Coefficient
PET_3s and # crashes	0.44
PET_2.5s and # crashes	0.45
PET_2s and # crashes	0.50
PET_1.5s and # crashes	0.62
PET_1s and # crashes	0.81

Table 5.17 shows the results from Fischer's test with respect to the PET parameters to determine their potential to identify high crash and low crash intersections respectively. It shows the results of the test on PET variables at four different thresholds of 3 sec, 2 sec, 1.5 sec and 1 sec. Both PET variables at thresholds of 1.5s, and 1 sec show the ability to classify intersections into the three categories. However, as the threshold increases, the variables PET_2s and PET_3s still show an ability to filter out low crash intersections but their CDF counterparts do not.

Table 5.17: Results of Fisher's exact test using PET parameters

Parameter	p-value from Fischer's test	
	High-Medium vs. Low	High vs. Medium-Low
PET_1s	0.23	0.50
PET_1.5s	0.23	0.23
CDF1s	0.23	0.50
CDF1.5s	0.23	0.50
PET_2s	0.23	0.09
PET_3s	0.23	0.04
CDF2s	0.04	0.09
CDF3s	0.04	0.09

Figure 5.7 shows the box plots for the rank of the parameter for intersections belonging to each category. Figure 5.8 shows the scatter plots of crash category of intersection with respect to the rank of the parameter. These Figures show that both parameters perform equally well for identifying high crash intersections. Figure 5.8 shows the actual performance of threshold of PET_1s and CDF1s. The red lines show that five of the six high ranked intersections are to the left of it and there are no other intersections in that cluster. This shows a clear demarcation by the red line at a rank of 5 for both parameters which convert into values of 6 and 1.92 respectively for PET_1s and CDF1s.

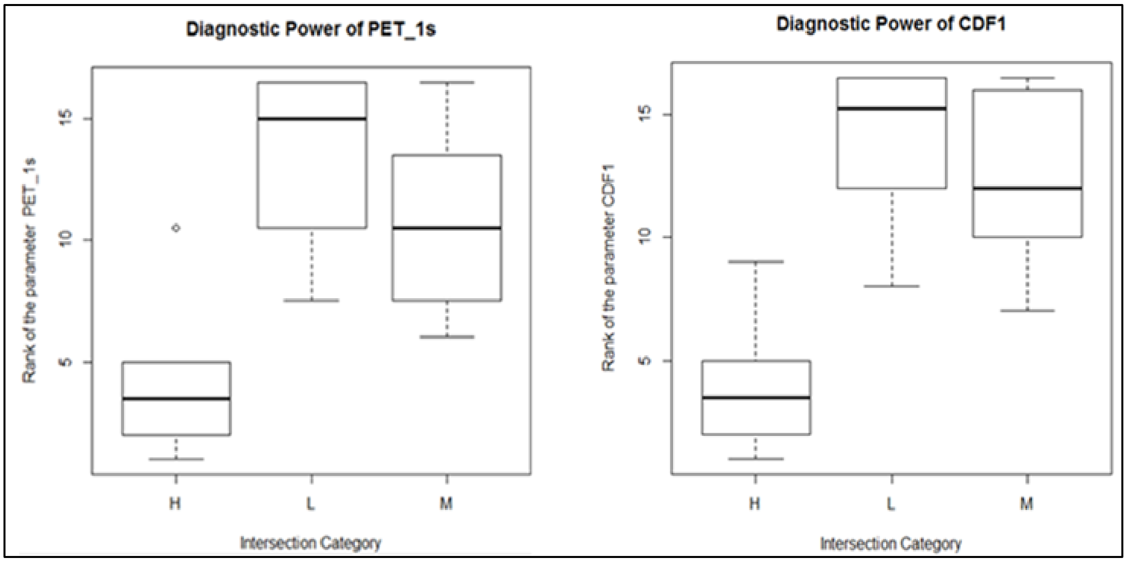


Figure 5.7: Box plots for evaluating the diagnostic power of PET measures

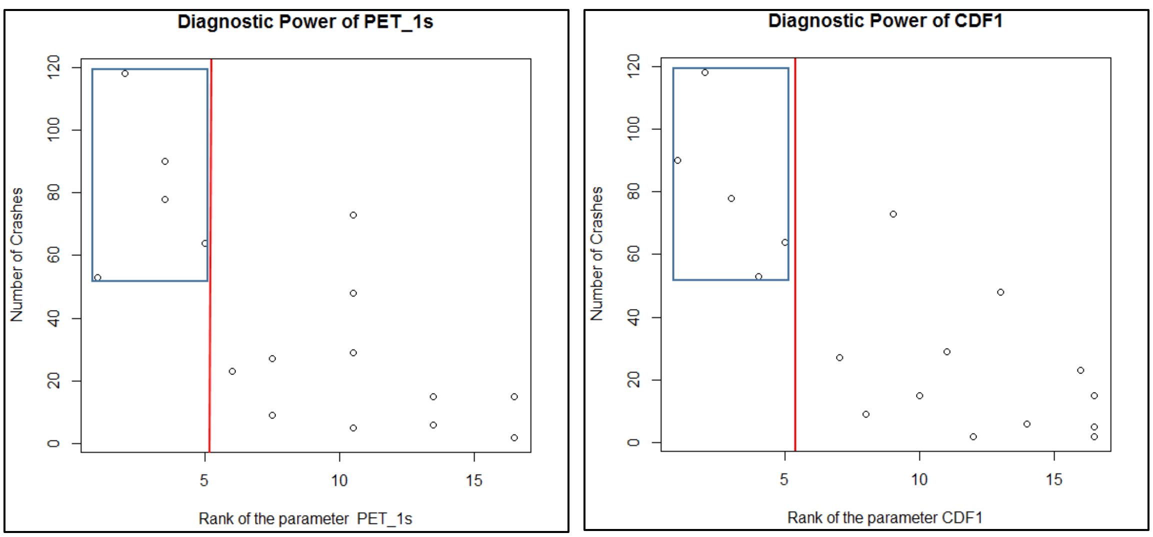


Figure 5.8: Scatter plots for evaluating the diagnostic power of PET measures

AADR values as shown in Table 5.18, support the conclusion that a threshold of 1 sec is much better than higher thresholds for identifying high crash intersections, but higher thresholds can be used to identify medium-low crash intersections. Given that AADR values with respect to low crash intersections are almost similar for PET thresholds greater than 1 sec, the parameter PET_1.5s, having a low AADR value for medium category intersections shows that it might be a better parameter to filter out medium and low crash intersections, provided sufficient data at 1.5s threshold is available. The scatter plots in Figure 5.9 and the box plots in Figure 5.10 also show that PET_15s looks like a better measure to identify low crash intersections (the box plots follow a logical order among the high, medium and low crash intersections). This shows that overall there is better confidence in the threshold and performance of PET_15s than in PET_3s.

Table 5.18: Identifying high and low crash locations using PET measures

Parameter	AADR-High Category Intersections	AADR-Low Category Intersections	AADR-Medium Category Intersections
PET_1s	2.2	2	3.1
PET_1.5s	3.5	3.2	1.8
CDF1s	1.5	1.9	3.4
CDF1.5s	2.8	3.3	2.5
PET_2s	5.8	3.3	3.7
PET_3s	6.3	3.3	2.5

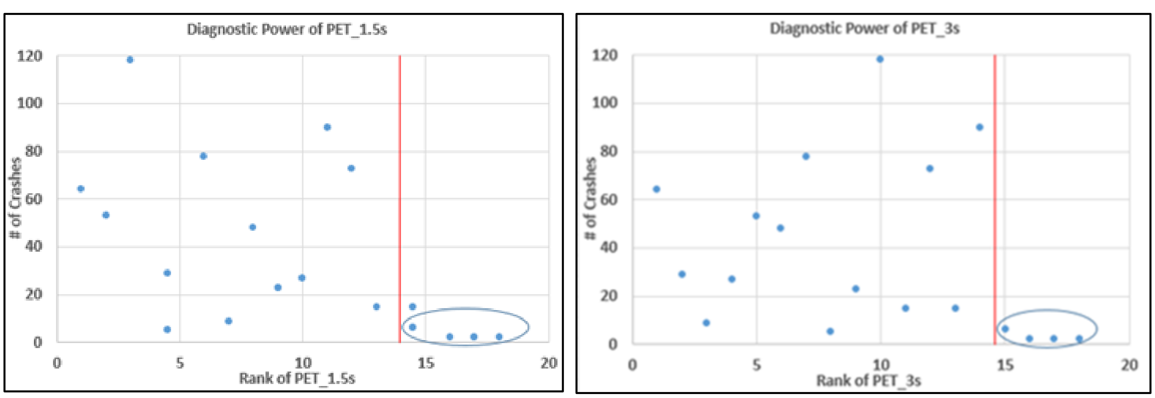
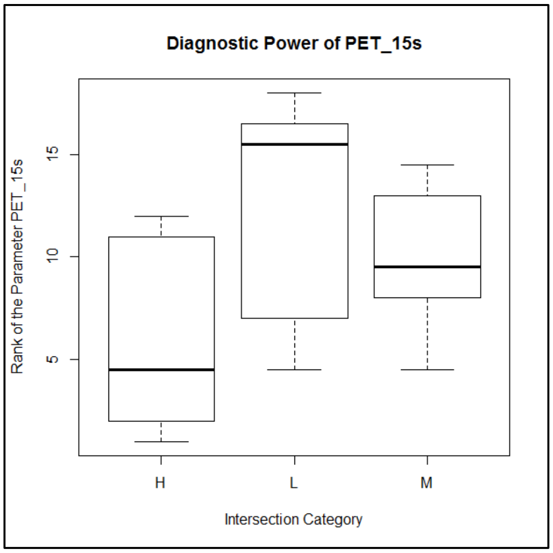


Figure 5.9: Scatter plots for evaluating the diagnostic power of PET measures



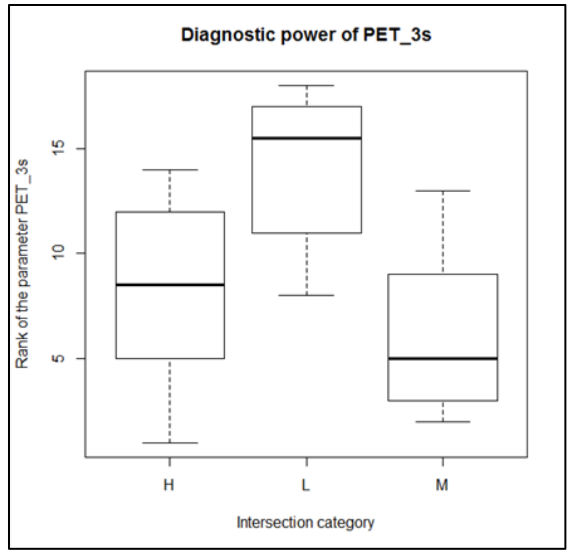


Figure 5.10: Box plots for evaluating the diagnostic power of PET measures

Figure 5.11 shows the scatter plots between the actual crash values and the PET parameter values instead of ranks to demonstrate the power of PET to group intersections, and the potential to identify parameter values that can act as thresholds to filter a category of intersections. The scatter plots clearly show that a threshold of 1 second has an ability to filter out high crash intersections and number of PETs below 1 second has the ability to filter out low crash intersections too. The green lines indicate potential threshold values of parameter for identifying category of an intersection. For example, based on the plots, values of 2 and 6 of the parameter number of PETs below 1 second can be used to filter out low crash and high crash intersections respectively, and a value of 1.92 for the parameter “proportion of PETs below 1 second (%)” to filter out high crash intersections. At 1.5 second threshold, “number of PETs” shows ability to filter out low crash intersections while “proportion of PETs” holds some strength to identify high crash intersections. However, the groupings for low and high crashes are very close to overlapping each other.

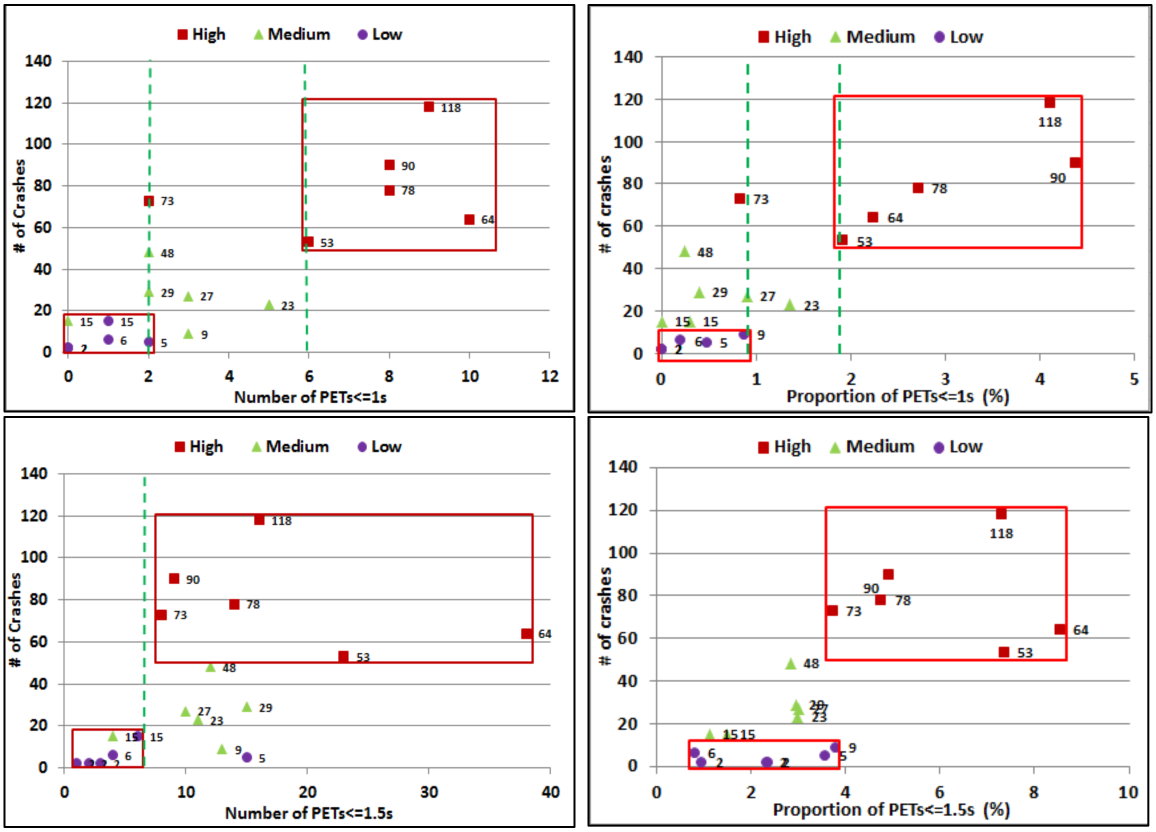


Figure 5.11: Scatter plots for between actual values of PET measures and crash numbers

5.5.3 Mixed Factor Analysis (PET + Volume)

One of the measures suggested in previous section was CDF value at various PET thresholds. The CDF value is computed as a proportion of total number of PETs observed. The total number of PETs observed is limited to the total number of permitted left-turn movements which is only a portion of the exposure term. Overall exposure is obtained by a combination of left-turn and through volumes which is termed "conflicting volume". The first parameter in the mixed factor category is the number of PETs below a threshold recorded as a proportion of total conflicting volume. Let us call this measure

PET rate. The correlation between PET rates at different thresholds and corresponding crash frequencies is shown in Table 5.19. Similar to the trend seen in other parameters, PET rate at 1 second threshold shows the highest RCC value with crash frequencies with a coefficient of 0.75 which is almost as high as the corresponding value observed for the parameters # of PETs below a threshold or CDF at a threshold.

Table 5.19: Rank correlation coefficients for the number of PETs as a proportion of total conflicting volume

Parameter	Rank Correlation Coefficient
PET_3s/Tot.Conf.Vol. and # crashes	0.08
PET_2.5s/Tot.Conf.Vol. and # crashes	0.18
PET_2s/Tot.Conf.Vol. and # crashes	0.27
PET_1.5s/Tot.Conf.Vol. and # crashes	0.51
PET_1s/Tot.Conf.Vol. and # crashes	0.75

Coming to the case of identifying high and low crash intersections, Table 5.20 shows that both parameters PET_1s/Conf. Vol. and PET_15s/Conf_Vol show the ability to filter low crash intersections. However, PET_1s/Conf_Vol only shows the ability to filter high crash locations. Table 5.20 shows that for identifying high crash intersections, the parameter PET_1s/Conf.Vol. has similar AADR value as found for simple PET parameters. However, an AADR value of 2.3 is still higher than the corresponding value for CDF1s parameter. For identifying low crash intersections, both the parameters PET_1s/Conf.Vol. and PET_1.5s/Conf.Vol. have almost equal AADR values which agrees with the analysis in section 5.5.2 where PET_1.5s was found to be effective in identifying low crash intersections.

Table 5.20: Identifying high crash locations using PET rate measures

Parameter	p-value from Fischer's test	
	High-Medium vs. Low	High vs. Medium-Low
PET_1s/Conf. Vol.	0.5	0.5
PET_15s/Conf. Vol.	0.5	0.09

Table 5.21: Identifying high and low crash locations using PET rate measures

Parameter	AADR-High Category Intersections	AADR-Low Category Intersections	AADR-Medium Category Intersections
PET_1s/Conf.Vol.	2.3	1.8	3.3
PET_1.5s/Conf.Vol.	5.2	2.7	3.0

Conflicting volume can also be taken into account by considering crash rates instead of frequency of crashes. Proportion of PETs below a threshold can be considered as normalization with respect to exposure. Hence it is pertinent to normalize crashes, and crash rate is one such measure normalized by exposure.

Exposure can be represented through two measures: AADT, and conflicting volume. It is established in the previous sections that a PET threshold of 1 second has the best total correlation with crashes in any form of the parameter. Hence Table 5.22 shows only RCC values at a threshold of 1 second. The correlations between the rankings of measures of PETs observed below a threshold of 1 second and corresponding crash rate measures (calculated based on AADT and conflicting volume) are shown in Table 5.22. These values show that all four combinations have almost equal rank correlations approximately equal to 0.8 which is similar to that found with other PET parameters with 1s as threshold. Hence both the measures (PET_1s and CDF1s) can be used to rank intersections in terms of crash rates too. However, it is important to recall that intersections having a major AADT of greater than 20,000 are considered in this study.

Table 5.22: Rank correlation coefficients between PET measures and crash rate

Parameter	Rank Correlation Coefficient
PET_1s and Crash Rate (AADT)	0.82
CDF1s and Crash Rate (AADT)	0.78
PET_1s and Crash Rate (Conf. Vol.)	0.79
CDF1s and Crash Rate (Conf. Vol.)	0.80

Coming to the case of identifying high and low crash rate intersections, the categories of intersections based on crash rates is not known as the categories were identified only based on crash frequency. Therefore classification tests cannot be performed in this case. However, AADR analysis is performed and the results are shown in Table 5.23. The Table shows that the AADR values generally are similar to those found for other PET measures. However, unlike other PET measures at a threshold of 3 sec, these parameters do not show an ability to filter out low crash rate intersections, given the higher AADR values corresponding to low and medium category intersections.

Table 5.23: Identifying high and low crash rate locations using PET measures

Parameter	AADR - High Category Intersections	AADR - Low Category Intersections	AADR - Medium Category Intersections
PET_1s and Crash Rate (AADT)	2.3	2.2	3.9
CDF1s and Crash Rate (AADT)	2.5	1.8	4.4
PET_1s and Crash Rate (Conf. Vol.)	2.6	2.2	3.3
CDF1s and Crash Rate (Conf. Vol.)	2.2	1.8	4.4
PET_3s and Crash Rate (AADT)	5.3	3.0	4.3
PET_3s and Crash Rate (Conf. Vol.)	6.3	3.0	4.3

5.6 SENSITIVITY ANALYSIS

Section 5.1 explained the criteria for the selection of intersections, and the reasons behind choosing the thresholds for “high”, “medium”, and “low” crash intersections. However, these thresholds are still subjective and can lead to a bias in the conclusions. This section discusses the analysis to explore the impacts of choosing different thresholds for these categories and if the conclusions drawn from the preceding sections still hold true. The non-parametric data analysis conducted in the preceding sections (using Fisher’s exact test) will be performed in this section too.

The section 5.5 concludes that the PET parameters at 1s are good at filtering out both high and low crash intersections, and that higher thresholds perform better in filtering out low crash intersections (1.5s being better than 3s). For evaluating the impact of intersection categorization on these conclusions, this chapter analyzes these parameters – PET_1s, CDF_1s, PET_1.5s, CDF_1.5s, PET_3s, and CDF_3s.

Table 5.24 shows the results of Fischer’s exact test to determine the performance of each of the parameters in filtering out “high” crash intersections, with different opposing left-turn crash frequency thresholds for classifying an intersection as “high” crash category. The Table lists the p-values from the test for each combination of crash threshold and parameter. The p-values marked red are cases where the null hypothesis is rejected. The Table shows that PET parameters at 1 second threshold are powerful in filtering out high crash intersections at any crash frequency level. As the threshold increases, it can be seen that the ability of the parameter to filter out high crash intersections decreases (Fischer's test gives an exact p-value. So, the p-values may be compared to determine relative performance). CDF_15s seems to be performing better then PET_15s (because it shows ability to categorize intersections having greater than 30 opposing left-turn crashes better than PET_15s. Moving the threshold higher to 3 seconds, the parameter PET_3s shows the ability only to differentiate between intersections on either side of 10 opposing left-turn crashes (PET_3s) while CDF_3s shows fails in all categories. The parameter CDF_1.5s on the other hand has shown the ability to even classify intersections having greater than 50 opposing crashes as "high" but fails at the 70 crash threshold (four of the study intersections have greater than 70 opposing left-turn crashes).

Table 5.24: Results of Fischer's Exact Test for combinations of threshold for "high" crash category

	Parameter					
Crash threshold for "High" category	PET_1s	CDF_1s	PET_15s	CDF_15s	PET_3s	CDF_3s
>=5	0.99	0.99	0.99	0.5	0.99	0.09
>=10	0.22	0.23	0.23	0.23	0.23	0.03
>=20	0.99	0.23	0.5	0.23	0.09	0.03
>=30	0.5	0.23	0.09	0.5	0.09	0.04
>=50	0.5	0.5	0.23	0.5	0.04	0.09
>=70	0.5	0.09	0.09	0.09	0.04	0.09

For identifying low crash intersections, Table 5.25 shows that the "number of PETs below a threshold" parameter seems to be performing better than their CDF counterparts. As there are no intersections with crash numbers below 2, nothing can be said about the performance of these parameters when intersections having 0 or 1 crash are considered to be belonging to “low” category. Higher threshold 3 seconds works as good as a threshold of 1 second when a “low” category is considered to be below 10 crashes, while a 1 second threshold shows ability to filter out “low” crash intersections, at any level of crash categorization. However, given the fact that the distribution of crashes is so heavily skewed towards low crash numbers, it is unlikely that a practical threshold for “low” category would be greater than 10 crashes.

Table 5.25: Results of Fischer’s Exact Test for combinations of threshold for “low” crash category

	Parameter					
Crash threshold for "Low" category	PET_1s	CDF_1s	PET_15s	CDF_15s	PET_3s	CDF_3s
<=5	0.99	0.99	0.99	0.5	0.99	0.23
<=10	0.5	0.5	0.23	0.23	0.5	0.05
<=20	0.5	0.23	0.23	0.09	0.09	0.04
<=30	0.5	0.23	0.23	0.23	0.09	0.04
<=50	0.23	0.23	0.09	0.5	0.09	0.03

The analysis presented above shows that as the threshold decreases, the power to classify very high crash intersections (having greater than 20 opposing left-turn crashes) increases. In terms of real world applicability, the optimal threshold selection for PET really depends on the level of classification (based on crashes) sought after. For example, if a transportation agency is required to classify intersections having greater than 30 crashes as a separate category to invest time and resources, then consideration for lower thresholds of PET (1.5 seconds) is required. If the requirement is to filter out the least safe intersections in terms of opposing left-turn crashes (four of the study intersections in this research are having greater than 70 opposing left-turn crashes in 4 years), then only a 1 second threshold has the power to identify such intersections.

5.7 SUMMARY

This chapter begins with a general description of the phase 3 of the research, the selection criteria for the study intersections, and explains the categorization of the study intersections into high, medium, and low crash categories based on the distribution of opposing left-turn crash numbers across intersections in the study area (Atlanta metro area).

The chapter then discusses the potential application of Generalized Extreme Value modeling technique to predict crash frequencies (or probability of crash) at an intersection just based on the PET data collected based on the fact that a PET value of 0 means crash. First of all, divergence between expected CDF values from GEV fit and observed CDF values are considerable for 1 second and 0 sec (crash) when any threshold value 2 seconds and higher is considered to fit the GEV distribution. The predicted crash probabilities from the GEV fits at thresholds higher than 1.5 second are more than the observed probabilities or in other words, crashes are over estimated. However, the difference between observed crash probability and predicted crash probability decreases significantly as threshold value decreases to 1.5. One possible explanation is that this means that the process of PET occurrence of higher PETs (especially 2 seconds and more) is different from that of 1.5 seconds and below (likely due to driver intervention to reduce the possibility of a crash). Hence thresholds greater than 1.5 s are not fit to predict crashes. This trend is seen in the data collected at other study intersections too.

At a threshold of 1.5s, on the other hand, there are only two data points (1.5 second and 1 second) which limits concluding about the data's ability to fit PET=0. Moreover due to unavailability of sufficient PET data below 1 second, it is not possible to determine the distribution at a threshold of PET=1 second and below.

The chapter then discusses the non-parametric rank-based data analysis conducted to determine the diagnostic power of PET to categorize intersections into high, medium, and low categories. Overall, considering all the 18 study intersections together to find rank correlation between PET and crashes, it can be seen that the measures PET_1s (number of PET values below 1 second) and CDF1s (proportion of PET values below 1 second) show the best rank correlation with crashes. Either of these two parameters could be used to rank intersections based on crash numbers. For the identification of high crash intersections (greater than 50 opposing left-turn crashes), a threshold of 1 second seems to be the best indicator (both PET_1s and CDF1s performing almost equally good). However, for identifying low crash intersections, firstly the parameter of "absolute number of PETs below a threshold" is a better indicator than CDF parameters. Second, it is not required to consider a threshold of as low as 1s because higher threshold of even 3s seem to be working well for filtering low crash intersections (given that "low" is defined as intersection having lesser than or equal to 10 opposing left-turn crashes). If sufficient data at 1.5 second threshold is available, PET_1.5s is a better parameter to filter out low crash intersections than the parameter PET_3s. Sensitivity analysis has shown that higher thresholds can also be used to filter out high-crash intersections if the threshold for classifying an intersection as "high" is at or below 50 opposing left-turn crashes. The next chapter explores the predictive power of PET as opposed to this chapter that explored the diagnostic power of PET.

CHAPTER 6: PARAMETRIC MODELING

Chapter 5 dealt with the effectiveness of PET parameters using non-parametric method by considering agreement between the ranks of PET values and crashes. The non-parametric analysis evaluated the diagnostic power of PET. It showed that the ability of PET to differentiate between high and low crash intersections varies with its threshold value. This chapter deals with the evaluation of predictive power of PET. This chapter begins with a basic analysis to determine if traffic volume, intersection safety characteristics, and PET have parametric correlation with crashes and how the correlation values are different from rank correlation values observed in chapter 5. The analysis is again divided into traffic volume based analysis, PET based analysis and mixed-factor analysis.

6.1 GENERALIZED LINEAR MODELING

6.1.1 Volume Based Analysis

The concept of exposure and its importance in safety research has already been explained in section 5.5.1 of chapter 5. Traffic volume is one of the most frequently used factors of safety representing exposure, and intuition suggests that the greater the exposure, the greater the likelihood for a crash. Current research deals with the number of vehicles entering the intersection as a preliminary measure of exposure which is represented by the measure AADT on major and minor roads of the intersection. The concept of conflicting volume as a more elementary unit of exposure has also been demonstrated in the section 5.5.1. Figure 6.1, Figure 6.2, and Figure 6.3 show the relationship between AADT measures and corresponding opposing left-turn crashes. Each of these plots has two regression fits (one is linear and the other is a 2nd degree polynomial fit). It can be seen that the polynomial fit has a better R^2 than the linear fit (in the cases of minor AADT and total AADT) although in all cases the R^2 value is generally low. These plots show that the maximum AADT does not necessarily correspond to maximum crashes. The likely reason is that as after some threshold value of AADT, the opportunities for making left turns decrease which in turn leads to lower opposing left-turn crashes.

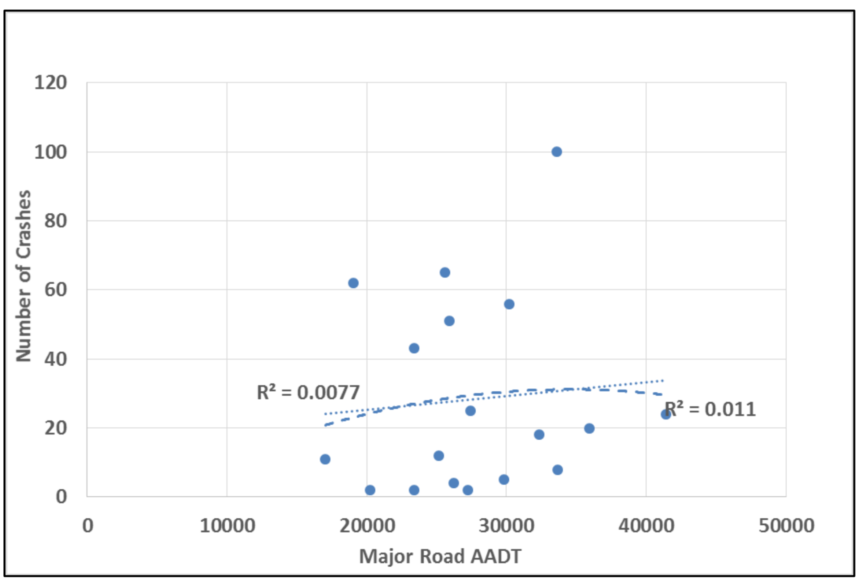


Figure 6.1: Relationship between major AADT and crashes on major roads

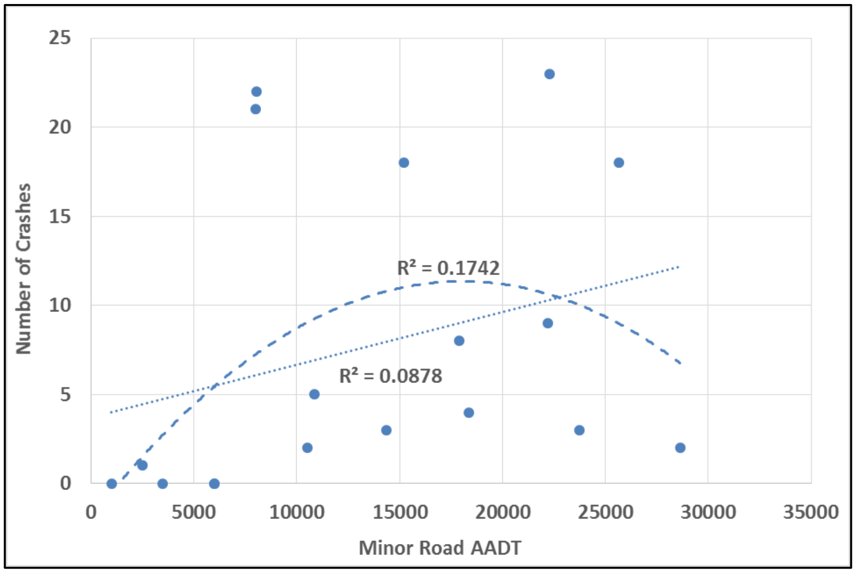


Figure 6.2: Relationship between minor AADT and crashes on minor roads

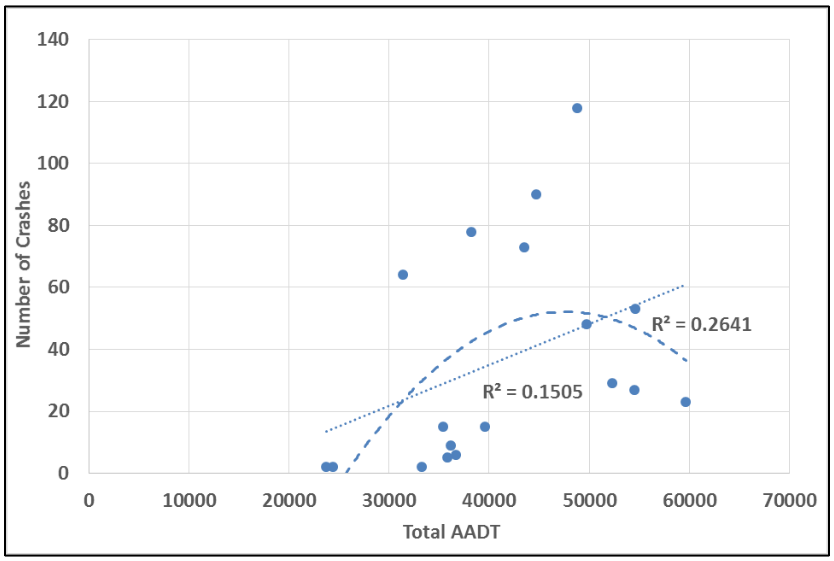


Figure 6.3: Relationship between total AADT and total crashes

Figure 6.4, Figure 6.5, and Figure 6.6 show the relationship between conflicting traffic volumes and crashes at these intersections. Even here, in the case of minor road, 2nd degree polynomial curve better fits the data than linear. However, the plots corresponding to major road and total conflicting volumes do not show any meaningful difference in fit between linear and 2nd degree polynomial. The most likely reason for this observation is that practically, as the conflicting volume increases, the signal is unlikely to have a permitted phase, and this created a limitation on the higher end of the conflicting volume that can be considered in our dataset, of which only intersections with a permitted phase were selected. In a world where there is no such bias, we would expect to see similar tend as a minor conflicting volume where there would be a downward trend in number of opposing left-turn crashed with increase in conflicting volume, as the opportunities to make a left-turn would decrease.

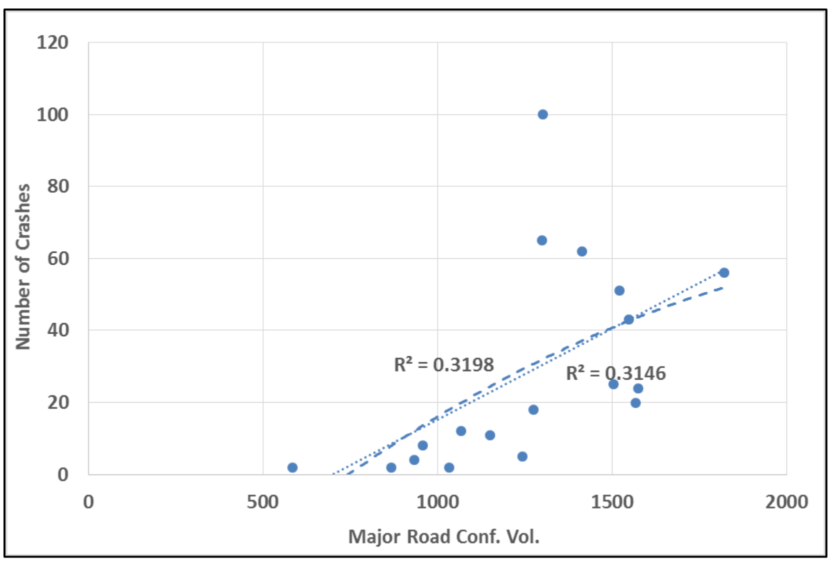


Figure 6.4: Relationship between major conflicting volume and major road crashes

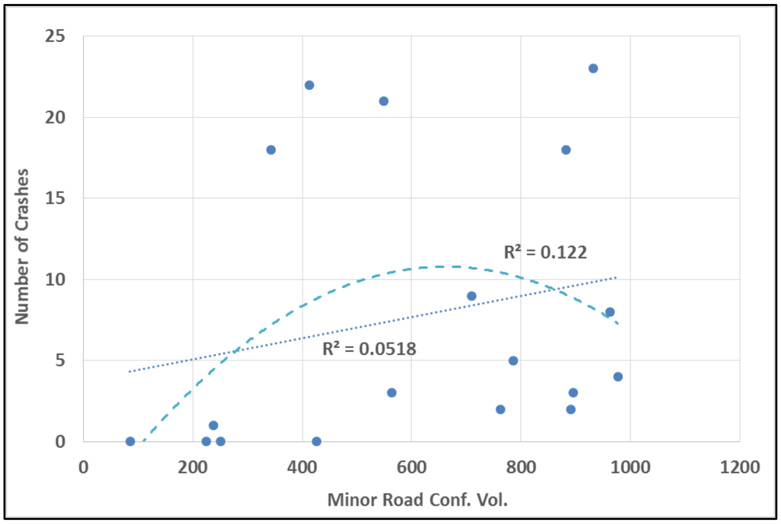


Figure 6.5: Relationship between minor conflicting volume and minor road crashes

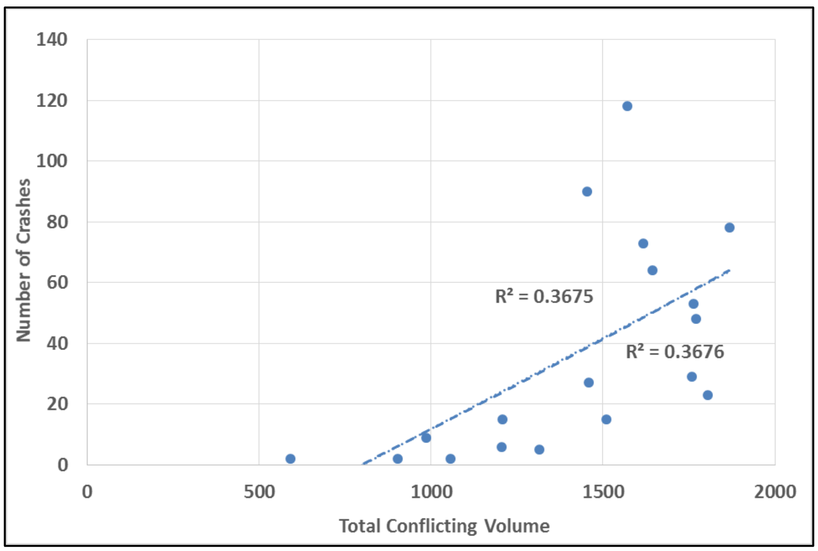


Figure 6.6: Relationship between total conflicting volume and total crashes

In this study, only intersections having a major road AADT greater than 20,000 were considered. This means that the range of traffic volumes and resulting conflicting volumes are restricted which might have resulted in lower R-squared values.

The analysis until now studied simple linear regression to estimate the relationship between traffic volume, and crashes. Linear regression uses ordinary least squares approach and makes a good first cut at the data. To some extent, it acts as a base model and guides the logical next step. However, standard linear regression models make certain assumptions which may limit its application to predict crashes. The first assumption is that the mean of response variable is a linear combination of predictor variables and that the response variable is normally distributed with this mean and constant variance. Linear models also have an underlying Gauss-Markov assumption that requires the error to be independent with a mean value of zero. Homoscedasticity is another major assumption of linear regression which means that the variance of errors in the response variables is constant (McCullagh and Nelder (1989)). Moreover, linear regression can predict negative, non-integer values depending on the regression equation developed. But crashes are positive whole numbers and linear regression might not be the most appropriate approach to model crashes.

Generalized linear models (GLM) stem from the concept that linear models can be transformed to create a framework that closely resembles linear models but can accommodate a wide variety of non-normal outcome variables. Nelder and Wedderburn (1972) presents one of the first attempts at developing this framework. A GLM consists of three major components.

- Random component: This specifies the characteristic distribution of the response variable with respect to the predictors. For example, this tells if the response variable follows a Gaussian, Poisson, Gamma or Negative Binomial distribution among others. Each of these forms is used to model certain types of data. Poisson distribution is often used to model count data. The Gaussian distribution is used to model data that follows a bell-shaped curve and is commonly known as the normal distribution. Gamma distributions are used to model continuous positive data.

- A linear predictor that is a linear function (η) of predictor variables on which the expected value of response (μ) depends.

  $$\eta = \alpha + \beta_1 X_1 + \ldots\ldots + \beta_n X_n \tag{6.1}$$

- A link function $g(\mu) = \eta$ which links the linear predictor of predictor variables to the expected value of response μ; g() should actually be called an inverse link function.

GLM retain their linear character through this link function by which the response and predictor are related. Because the linear predictor is a linear function of explanatory variables, the linear assumption is preserved. However, it should be noted that retaining the linear component can be a limitation of this approach as well. Moreover, the distributions are restricted to certain families (e.g. exponential) and responses are constrained to be independent. Generalized Linear Modeling is not applicable to situations where predictor variables are auto-correlated or responses follow a time-series pattern unless additional steps are included.

While there are several distributions that can be used, the most commonly used methods to model crash counts use Poisson and Negative Binomial regression (Hauer et al., (1988)), as crashes have a very small probability of occurrence and they can be classified as count data. The following section describes in detail these regression approaches.

This part of the analysis was performed using R software package (Venables et al. (2013)). The function "glm" in R is specifically used to perform the generalized linear modeling analysis.

6.1.2. Poisson Regression (Hauer et al., (1988))

Poisson regression assumes that the observed counts are generated from Poisson distribution. The Poisson distribution is often used to model count data and events that have a very small probability of occurrence, e.g. telephone calls arriving in a system, vehicles arriving at a traffic signal, number of claim applications coming to an insurance company etc. The probability mass function of a Poisson distribution is

$$P(Y=y) = \lambda^y e^{-\lambda}/y! \tag{6.2}$$

where

λ = mean number of events in a unit time

y = value of the random variable for which the probability is being estimated

The relation between GLM and Poisson regression is that the mean of the Poisson distribution λ is estimated from the linear predictor of explanatory variables using the link function. The most common link function is the log link which is expressed as

$$\log(\lambda) = \eta = \alpha + \beta_1 X_1 + \ldots\ldots + \beta_n X_n$$

$\Rightarrow \lambda = \exp(\alpha + \beta_1 X_1 + \ldots\ldots + \beta_n X_n)$ (6.3)

where

$X_1, \ldots, X_n$ are the explanatory variables and $\beta_1, \ldots\ldots, \beta_n$ are regression coefficients.

6.1.3 Negative Binomial Regression (Hauer et al., (1988))

One of the major properties of a Poisson process is that the mean of the distribution is equal to the variance. Often times, especially in dealing with crash counts, this property is violated. Data is said to be under-dispersed if variance is less than mean, and over-dispersed if variance is greater than mean. Negative binomial regression is normally used in the case of over-dispersed data. Suppose that $Y \sim \text{Poisson}(\lambda)$ and that λ itself is a random variable with a Gamma distribution i.e., $\lambda \sim \text{Gamma}(\alpha, \beta)$ with mean $\alpha\beta$ and variance $\alpha\beta^2$. The probability density function of this distribution is:

$f(\lambda) = (1/\beta^{\alpha}\,\Gamma(\alpha))\,\lambda^{\alpha-1}\exp(-\lambda/\beta)$ (6.4)

It can be shown that in such a case, Y follows a negative binomial distribution:

$f(Y) = [\Gamma(\alpha+y)/(\Gamma(\alpha)y!)]\,[\beta/(1+\beta)]^{y}\,[1/(1+\beta)]^{\alpha}$ (6.5)

This distribution has a mean $\alpha\beta$ and variance $\alpha\beta + \alpha\beta^2$. The negative binomial model is generally expressed in terms of parameters $\mu = \alpha\beta$ and an overdispersion parameter $K = 1/\alpha$. This makes

$E(Y) = \mu$ and $Var(Y) = \mu + K\mu^2$. (6.6)

In terms of the parameters μ and K,

$f(Y) = [\Gamma(1/K+y)/(\Gamma(1/K)y!)]\,[K\mu/(1+K\mu)]^{y}\,[1/(1+K\mu)]^{(1/K)}$ (6.7)

It can be seen that as K -> 0, the distribution approaches Poisson distribution with mean μ. In terms of GLM regression,

$Y \sim NegBin(\mu,K)$

Assuming a log link,

$$\log(\mu) = \alpha + \beta_1 X_1 + \ldots\ldots + \beta_n X_n \tag{6.8}$$

Table 6.1 and Table 6.2 show values of some of the intersection characteristics and variations of surrogate measure (PET) considered in the GLM analysis presented below. Although many other intersection characteristics were considered in the GLM analysis, only those that consistently appear in the analysis presented are shown here. All other variables and models developed using them can be found in the appendix where scatter plots of the relationship between actual number of crashes at the study intersections and some of these factors are also presented.

Table 6.1 PET measures considered in the GLM analysis

Intersection	Crashes	PET_15s	PET_1s	PET_15sbyPET_3s	PET_1sbyPET_3s	CDF1	CDF15
N Druid Hills Rd and Lavista Rd	118	16	9	26.2	14.8	4.11	7.31
Roswell Rd and W Wieuca Rd	78	14	8	19.4	11.1	2.72	4.76
Lawrenceville Hwy and Lawrenceville Suwanee Rd	73	8	2	18.4	4.1	0.83	3.73
GA 138 and Sigman Rd	90	9	8	22.5	20	4.37	4.92
GA 20 and Willow Lane	64	38	10	31.4	8.3	2.25	8.54
Grayson Hwy and Scenic Hwy	53	23	6	27.7	7.2	1.92	7.37
N Druid Hills Rd and Lawrenceville Hwy	48	12	2	16.2	2.7	0.24	2.84
GA 10 and Grayson Pkwy	29	15	2	14.2	1.9	0.39	2.95
Ponce De Leon Ave and Moreland Ave	27	10	3	10.6	3.2	0.91	3.02
Scott Blvd and Clairemont Ave	23	11	5	17.7	8.1	1.36	2.99
Memorial Dr and Covington Hwy	15	4	0	8.3	0	0.00	1.11
GA 10 and Oak Rd	9	13	3	13.7	3.2	0.87	3.79
Glenwood Dr and Columbia Dr	15	6	1	9.5	1.6	0.48	3.57
Sugarloaf Pkwy and Buford Hwy	6	4	1	18.2	4.5	0.20	0.81
MLK Jr Dr and Brownlee Rd	2	2	0	7.4	0	0.30	1.48
Cobb Pkwy and Gresham Rd	5	15	2	25	3.3	0.00	2.33
Whitlock Ave and LindleyAve	2	1	0	3.8	0	0.00	2.37
North Ave and Techwood Dr	2	2	0	10.5	0	0.00	0.97

Table 6.2 Intersection characteristics considered in the GLM analysis

Intersection	Crashes	Major_AADT	Minor_AADT	Sqrt_Prod	ConfTot	Max_Grade	Min_Ln	MinSD	DiffSD
N Druid Hills Rd and Lavista Rd	118	33600	15205	22603	1572	3.78	10.93	300	-100
Roswell Rd and W Wieuca Rd	78	30180	8020	15558	1868	1.52	8.6	300	-100
Lawrenceville Hwy and Lawrenceville Suwanee Rd	73	25594	17896	21402	1617	2.50	10.98	620	130
GA 138 and Sigman Rd	90	19060	25655	22113	1454	2.49	11.1	450	0
GA 20 and Willow Lane	64	23380	8000	13676	1643	2.62	9.43	800	310
Grayson Hwy and Scenic Hwy	53	25918	28629	27240	1762	2.05	11.9	500	100
N Druid Hills Rd and Lawrenceville Hwy	48	27415	22280	24714	1770	2.00	9.77	400	-40
GA 10 and Grayson Pkwy	29	41400	10853	21197	1759	2.60	11	600	110
Ponce De Leon Ave and Moreland Ave	27	32320	22220	26798	1460	2.00	8.4	420	20
Scott Blvd and Clairemont Ave	23	35925	23725	29195	1805	3.00	10	400	0
Memorial Dr and Covington Hwy	15	25180	14365	19019	1208	1.40	11.09	800	400
GA 10 and Oak Rd	9	33630	2500	9169	986	1.30	10.3	800	400
Glenwood Dr and Columbia Dr	15	17025	18360	17680	1509	2.67	8.71	480	-10
Sugarloaf Pkwy and Buford Hwy	6	26210	10500	16589	1205	1.55	10.18	800	310
MLK Jr Dr and Brownlee Rd	2	23360	1000	4833	592	4.10	10.63	520	30
Cobb Pkwy and Gresham Rd	5	29820	6000	14653	1314	3.14	10.05	440	-50
Whitlock Ave and LindleyAve	2	27230	6000	12782	904	1.15	10.28	800	440
North Ave and Techwood Dr	2	20240	3470	8381	1057	2.40	9.42	450	90

Where:

PET_15s = cumulative number of PETs below 1.5sec

PET_1s = cumulative number of PETs below 1sec

PET_15sbyPET_3s = fraction of PETs below 3sec that are below 1.5sec (in %)

PET_1sbyPET_3s = fraction of PETs below 3sec that are below 1sec (in %)

CDF1 = cumulative distribution function value of PET of 1 sec

CDF15 = cumulative distribution function value of PET of 1.5 sec

Max_grade = max average grade of the four approaches to the intersection

Min_ln = minimum width of approach lanes of an intersection

Sqrt_Prod = square root of the product of major road and minor road AADTs

Conf_Tot = sum of conflicting volumes of major and minor road

MinSD = minimum sight distance available for left-turn vehicles at an intersection

DiffSD = difference between the required sight distance and the minimum available sight distance

The first factor analyzed in the scatter plots is approach grade (Figure 6.7). The methodology to measure grade has been described previously in the “Methodology” chapter. For each approach to an intersection, a grade value was computed. Overall, three types of grade measures have been studied: average major approach grade, average minor approach grade, and maximum approach grade. Though many of the previous research works (Hauer (1992), Shankar (1994), Hauer (2010), Ahmed (2011)) show that grade is a factor in determining the safety of intersections, grade data from the 18 study intersections does not indicate any significant relationship with left-turn opposing crashes. Dividing the intersections into high, and medium-low categories in order to determine if the effect of grade is different on different categories of intersections, this plot shows some positive correlation of maximum approach grade with crashes for high intersection category while medium-low category intersections do not show any correlation. The study by Shankar (1994) shows that maximum grade greater than 2% increases rear-end crashes. Even the crash prediction models presented in HSM (2011) considers grade only for rural 2-lane highways. Since, the intersections being considered here were urban, and the crashes being considered are of the opposing left-turn type, grade was not a major determining criterion for the selection of intersections. However, every effort was made to select intersections having varied grade characteristics.

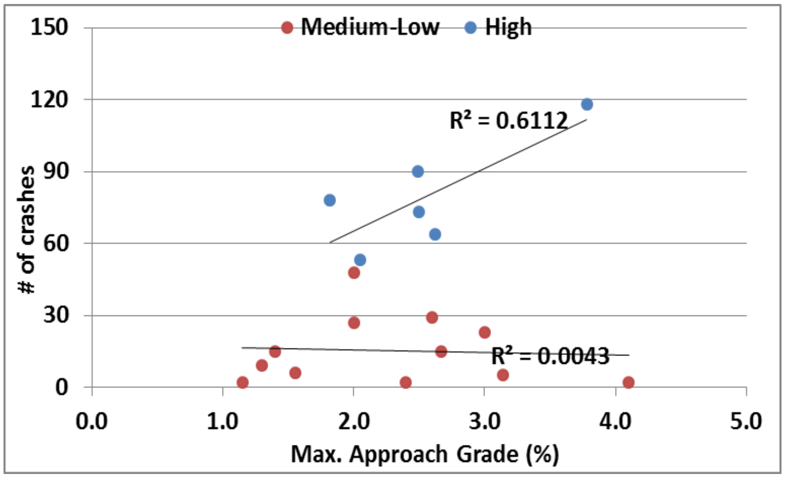


Figure 6.7: Scatter plot of maximum approach grade at an intersection

Figure 6.8 shows the relationship between crashes and minimum approach lane width. Approach lane width can have two conflicting types of effects on safety of left-turn vehicles. Smaller lane widths can have sight distance issues when left-turn vehicles in opposite lanes block the view of through vehicles for each other.

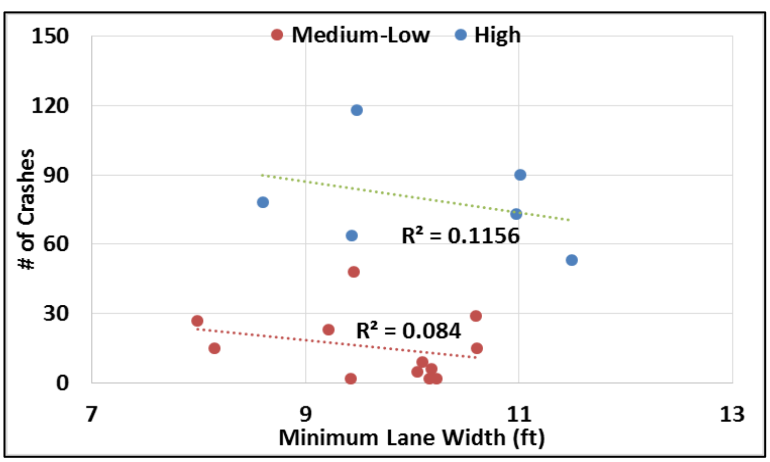


Figure 6.8: Scatter plot of minimum approach lane width

However, when lane width increases, left-turning vehicles need more time to completely cross the area of conflict and hence need more gap which decreases their safety. Again dividing the intersections into high, and medium-low categories, this plot shows some negative correlation (though of a small value) of minimum lane width with crashes for high intersection category while medium-low category intersections do not show any correlation.

The third factor presented here is sight distance (Figure 6.9 (a) and Figure 6.9 (b)). Sight distance is generally considered one of the most important factors of safety (Glennon (1987)) and intersections at which vehicles are permitted to turn left across opposing traffic should have sufficient sight distance to enable the drivers involved take safe decisions. In 2001, the American Association of State Highway and Transportation Officials (AASHTO) in its Policy on Geometric Design for Streets and Highways (2004) (Green Book) added standards for sight distance required for left-turn vehicles at intersections during permitted movements. It provides sight distance based on a left turn by a stopped vehicle, since a stopped vehicle requires more time to complete the turn than a vehicle that turns left without stopping. The sight distance along the opposing approach is the distance traversed by the through vehicle at the design speed in the time that a left turn vehicle requires to complete the turn (Table 6.3).

Table 6.3: Left turn gap required for computing sight distance (AASHTO (2004))

Design Vehicle	Time gap (t_g) (sec) at design speed of major road
Passenger Car	5.5
Single-unit Truck	6.5
Combination Truck	7.5

Speed limits of the approach roads at the study intersections are readily available from the field. However the sight distance computation requires design speed. The design speed for this analysis was assumed to be 10 mph greater than speed limit (Donnell et.al. (2009)). Field visits of the study intersections have shown the minimum left turn sight distance available at an intersection among the four approaches. For this approach of each intersection, the required intersection sight distance was calculated based on the assumed design speed. Figure 6.9 (a) and Figure 6.9 (b) show the relationship of two measures of sight distance with crashes.

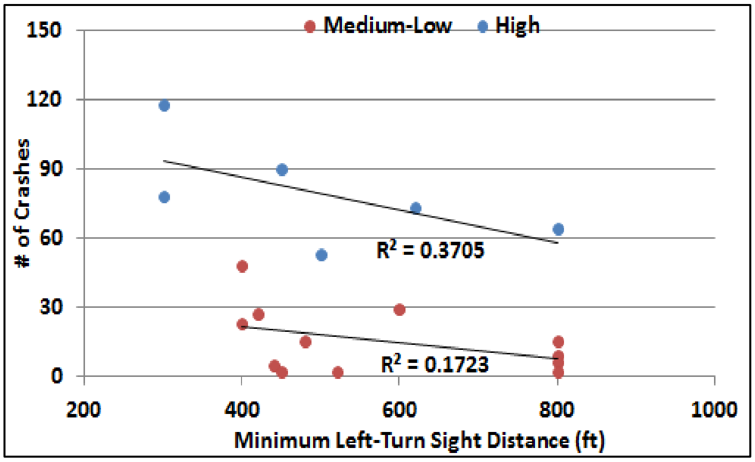


Figure 6.9 (a): Scatter plots of minimum available left turn sight distance and crashes

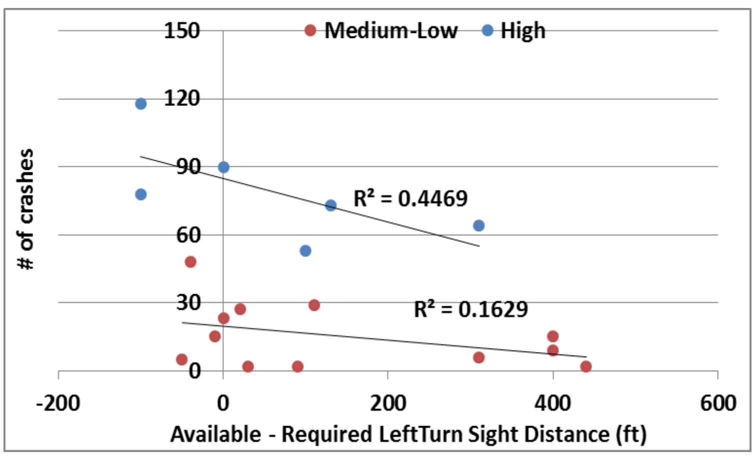


Figure 6.9 (b): Scatter plots of the crashes and the difference between the required left turn sight distance and available sight distance

It can be seen that minimum available sight distance has some limited correlation with crashes but the difference between available and required sight distance seems to have a better relationship. Dividing the intersections into high and medium-low groups (based on crash frequencies), it can be seen that both the variables- available sight distance, and difference between available and required sight distance are better correlated with crashes in case of high category compared to medium-low category and these correlations are better than all the intersections considered together. However, these coefficients are significantly lower than those found between crashes, traffic volume, and PET.

The observations found here are complementary to those found with traffic volume and conflicting volume (as shown in section 6.1.1) where the traffic volume parameters had good correlation with low and medium crash intersections in comparison to high crash intersections.

6.1.4 PET Based Models

It has already been established in Chapter 5 that PET threshold values of 2 seconds and above do not have good rank correlation with crashes. This means that those thresholds would likely not have good "predictive" correlation with crashes either. Thus, thresholds values of 2 and higher are not considered in the parametric modeling.

The first category of GLM models presented in this section are those that are driven only by PETs. The parameters are first estimated using Poisson regression. Table 6.4 shows the results for Poisson regression using PET measures as parameters. The Table shows different models using Poisson family using three different links – "log", "sqrt", and "identity".

The Akaike Information Criteria (AIC) shows the relative goodness-of-fit of models with respect to each other (McCullagh and Nelder (1989)) and it provides for a means of model selection. AIC does not tell how well the model fits the data on an absolute basis. It just provides for relative assessment among a set of models. It is a criterion that balances fitting the data with a model and number of parameters used in the model. It penalizes models having large number of predictors. A good model has a balance between bias and variance and hence AIC provides a measure to obtain that balance. The smaller the AIC, the relatively better the mode is.

$$AIC = -2*\ln(likelihood) + 2*k \tag{6.9}$$

Where:

ln = natural logarithm

likelihood = given the model and parameters, the probability of the observed outcomes from which the model is built

K = number of free parameters in the model

Table 6.4: PET-only based models using the Poisson family of distributions

Family (Link)	Variable	Estimate	P-value	Null Deviance	Residual Deviance	AIC
Poisson (log)	Intercept	2.666	< 2e-16			
	PET_1s	0.2012	< 2e-16	586.92	257.73	348.01
	Intercept	3.1209	< 2e-16			
	PET_15s	0.0371	< 2e-16	586.92	493.29	583.97
	Intercept	2.8617	< 2e-16			
	PET_15sbyPET_3s	0.1039	< 2e-16	586.92	278.65	368.93
	Intercept	2.1033	< 2e-16			
	PET_15sbyPET_3s	0.0789	< 2e-16	586.92	355.21	445.49
	Intercept	2.8414	< 2e-16			
	CDF1	0.454	< 2e-16	586.92	235.28	325.55
	Intercept	2.4484	< 2e-16			
	CDF15	0.2661	< 2e-16	586.92	308.34	398.62
Poisson (sqrt)	Intercept	3.2528	< 2e-16			
	CDF1	0.6868	< 2e-16	586.92	234	324.28
Poisson (identity)	Intercept	9.209	< 2e-16			
	CDF1	23.547	< 2e-16	586.92	187.96	278.24

The above Table does not show all the models tested. All the three link functions show that the parameter PET_1s is significant. For the log link, different forms of PET data are shown in the Table and it shows that PET_1s (absolute number of PETs below 1 sec) has the best fit among the models considered based on AIC values. Therefore, for the remaining two link functions presented, only models with PET_1s are shown. However, AIC only tells the relative goodness-of-fit. However, the AIC value does not establish the statistical significance of the model.

The results shown in the Table 6.4 also include the null deviance and residual deviance. The null deviance is the deviance associated with a null model or a model with no predictor variable and the deviation in such a case is calculated merely from the intercept term. All the models shown in Table 6.4 have a high null deviance, and indicate a very little fit (significant difference between the observed values and fitted values). The residual deviance on the other hand measures how much the data deviates from the current model. According to (McCullagh and Nelder (1989)), for a well-fitting model, the residual deviance should be approximately equal to the residual degrees of freedom (RDoF) and certainly lesser than about 1.5 times of the RDoF. When this is not the case, either the model does not adequately describe the variation in the data or the data are said to be exhibiting overdispersion.

For all the models considered in Table 6.4, the residual degrees of freedom is 16 and it can be clearly seen that the residual deviance values are much larger than the corresponding degrees of freedom. This shows that the crash data exhibits overdispersion and that Poisson regression model might not be the appropriate model for this. For most error structures, deviance is distributed asymptotically as chi-square (χ^2). So, the goodness-of-fit of a model can also be evaluated by testing the deviance against the chi-square distribution with appropriate degrees of freedom. The residual deviance of 163.97 when compared to a chi-square distribution with 16 degrees-of-freedom resulted in a p-value of lesser than 0.001. Therefore, there is sufficient evidence to reject the null hypothesis that Poisson regression is a good fit. This also implies that the crash numbers display overdispersion.

The second family of models studied in this section belongs to the Negative Binomial distribution and only the log link could be fit for the data. A Negative Binomial (NB) distribution is also called as Poisson-Gamma distribution as the NB distribution can be viewed as a Poisson(λ) distribution, where λ is itself a random variable, distributed according to a Gamma distribution. The negative binomial models are fit using the "glm.nb" command of MASS package for R. Table 6.5 shows the details of the models tested and it shows that with the NB family, the parameters PET_1s and CDF15 are significant at the 0.001 level while the remaining parameters are significant at 0.005 level. The AIC values in the Table show that the three parameters PET_1s, CDF1, and CDF15 have almost the same AIC values and any difference is statistically insignificant. Hence AIC does not show any significant best-fit model among the ones considered. The p-value for the chi-square test is 0.23 for all the considered models. Though this means that there is no sufficient evidence to reject the hypothesis that a negative binomial model is a good fit, the goodness-of-fit of the models using the considered parameters are not statistically different.

Table 6.5: PET-only based models developed using NB family

Family (Link)	Variable	Estimate	p-value	Dispersion	Null Deviance	Residual Deviance	AIC
NB (log)							
	Intercept	2.47	< 2e-16	1.71	36.842	19.639	159.62
	PET_1s	0.24	1.97E-05				
	Intercept	2.69	< 2e-16	1.6	34.579	19.743	160.92
	CDF1	0.55	9.30E-05				
	Intercept	2.61	5.40E-12	1.11	24.723	20.035	167.75
	PET_15s	0.08	0.00373				
	Intercept	2.01	1.47E-07	1.55	33.659	19.606	161.3
	CDF15	0.37	3.18E-05				

The mathematical equations for the negative binomial regression models with PET_1s parameter is as follows:

$$y = e^{(2.47+0.24*PET\_1s)} \tag{6.10}$$

The above analysis has shown that

- The hypothesis that a Poison error structure would be a good fit to model the crash data using PET could be rejected based on the result from a chi-square test.
- There was not sufficient evidence to reject the hypothesis that negative binomial error structure would be a good fit.

However, this analysis could not show any statistical difference in the goodness-of-fit using different parameters in the negative binomial model, except for very minor differences in the AIC and deviance values.

6.1.5 Combined Model

The last set of GLM models are developed using a combination of surrogate measure (PET) and intersection characteristics. Since combined models are being considered here, these models definitely have increased complexity compared to those which consider just surrogates or intersection characteristics. Table 6.6 shows the models developed using Poisson regression. It has already been seen that Poisson model is not an appropriate form to model crashes with only PET as parameter.

Table 6.6: Combined models using Poisson family

Family (Link)	Variable	Estimate	P-value	Null Deviance	Residual Deviance	AIC
Poisson (log)	Intercept	-1.633	0.00564			
	CDF1	0.305	< 2e-16			
	ConfTot	0.002	< 2e-16			
	min_ln	0.167	3.96E-05			
	Min_SD	-0.001	4.31E-13	586.918	36.761	133.04
	Intercept	-10.89	0.0152			
	PET_1s	0.124	< 2e-16			
	ConfTot	7.48E-03	0.00437			
	min_ln	1.121	0.00963			
	Min_SD	-2.18E-03	< 2e-16			
	ConfTot*min_ln	-5.08E-04	0.04607	586.918	52.088	150.37
Poisson (sqrt)	Intercept	-1.18E+01	1.58E-05			
	CDF1	1.186	< 2e-16			
	ConfTot	5.58E-01	2.36E-05			
	min_ln	8.59E-03	1.82E-09			
	Min_SD	6.58E-03	0.05166			
	ConfTot*Min_SD	-7.40E-06	0.00139	586.918	30.328	128.61
	Intercept	-1.36E+01	4.57E-07			
	PET_1s	4.83E-01	< 2e-16			
	ConfTot	8.59E-03	2.23E-09			
	min_ln	7.73E-01	1.94E-09			
	Min_SD	7.16E-03	0.034187			
	ConfTot*Min_SD	-8.56E-06	0.000211	586.918	53.415	151.69

Though the deviance values in the combined models are much lower than those observed before, they are still significantly high. A chi-square goodness-of-fit test of the model having the lowest AIC value results in a p-value of lesser than 0.001 and hence, in terms of statistical significance, the basic hypothesis that Poisson model is a good fit can be rejected even in this case.

The next set of models developed using a combination of surrogate measure (PET) and intersection characteristics use a Negative Binomial distribution. Table 6.7 and Table 6.8 show some of the models developed using the NB structure. Table 6.7 shows models only involving a combination of surrogate measure and traffic volume characteristic. Table 6.8 shows some models including more intersection characteristics.

A chi-square test of the above fit models in Table 6.7 shows that for all the above models, the p-value is greater than 0.2 which shows that the hypothesis that negative binomial form is a good fit for the above models cannot be rejected and that there is no statistical difference in the goodness-of-fit of the above models. Though the residual deviances in the combined models are lesser than models having only PET parameters, the reduction is marginal from 19.6 to 16.88. Given that a good fit GLM model will have residual deviance almost equal to degrees of freedom, a model containing CDF1 and TCV (conflicting volume) comes out to be the best-fit. The equation for the model is:

$$y = e^{(0.44354*CDF1+0.002475*TCV)} \tag{6.11}$$

Since negative binomial models are fit on data that express overdispersion (variance greater than mean), R-squared for a fit between observed and fitted crash values (as in the case of OLS) cannot be used to interpret the fit of the negative binomial models. Negative binomial regression does not have an equivalent to the R-squared measure found in OLS regression. MacFadden's adjusted R-squared (Smith and McKenna (2013)):

$$R_{adj}^2 = 1- (LLFM + K)/LLNM \tag{6.12}$$

Where,

R_{adj}^2 – Adjusted R-squared

LLFM – log-likelihood of full model

LLNM – log-likelihood of null model

K – number of parameters

For the negative binomial model specified in the equation, the adjusted R-squared can be computed as 0.802. However, this cannot be interpreted in comparison to R-squared from an OLS regression. Nevertheless, the adjusted R-squared tell that the model is a good fit. However, based on the principle of parsimony, it can be concluded that PET at 1 second threshold (PET_1s) (based on its lowest AIC value) shows the best fit and on its own has statistically significant predictive power.

Table 6.7: Combined models using PET and traffic characteristics

Family (Link)	Variable	Estimate	p-value	Dispersion	Null Deviance	Residual Deviance	AIC
NB (log)	Intercept	5.35E-01	0.25	3.13	62.378	17.594	149.92
	PET_1s	2.20E-01	1.47E-06				
	Sqrt_Prod	1.02E-04	1.48E-05				
	Intercept	-0.95791	0.107	6.19	107.8	16.88	139.68
	CDF1	0.44354	4.14E-08				
	TCV	0.002475	5.21E-10				
	Intercept	-1.04406	0.130461	3.22	63.91	17.73	149.62
	CDF15	0.237221	0.000957				
	TCV	0.002351	3.36E-06				
	Intercept	-2.20862	0.00493	9.74	149.702	16.773	136.04
	CDF1	2.291768	0.00226				
	TCV	0.003307	9.37E-11				
	CDF1*TCV	-0.00118	0.01217				

Table 6.8 shows results of negative binomial fit for models containing intersection characteristics in addition to traffic volume or conflicting volume. It is seen that intersection characteristics such as Min_SD (minimum sight distance), min_minor (minimum minor lane width) are some of the additional intersection characteristics that are significant. The lowest residual deviance (17.97) is observed for a model containing PET_1s, TCV, and Min_SD. However, the complexity of model has increased in comparison to models shown prior to this. The residual deviance is much higher than the degrees of freedom (14). Though the residual deviance of 17.97 shows that the model is a good fit (p-value of 0.25 in for a chi-square test with 14 degrees of freedom), it does not justify the additional complexity in the model. Therefore it can be said that a NB model with just the PET measure provides a good fit with significant predictive power, but inclusion of further intersection characteristics do not justify the increase in complexity.

Table 6.8: Combined models using PET, traffic Characteristics, and intersection characteristics

Family (Link)	Variable	Estimate	P-value	Null Deviance	Residual Deviance	AIC
NB (log)	Intercept	-15.69	0.000249			
	PET_1s	0.8298	0.032584			
	min_minor	0.3734	0.074995			
	log(sqrt_Prod)	1.56E+00	4.15E-06			
	PET_1s*min_minor	-0.0612	0.110086			
	Min_SD	-1.62E-03	0.010053	115.479	18.565	146.2
	Intercept	-3.10E+01	2.62E-11			
	CDF1	1.99E+01	1.95E-10			
	min_minor	3.11E-01	2.29E-13			
	log(sqrt_Prod)	3.17E+00	3.72E-04			
	CDF1*sqrt_Prod	-1.97E+00	8.15E-03			
	Min_SD	-1.10E-03	4.60E-10	256.505	19.484	133.82
	Intercept	-1.61E+01	8.25E-05			
	PET_1s	1.57E-01	1.80E-05			
	Min_SD	-1.93E-03	1.97E-03			
	TCV	2.73E+00	5.70E-01	116.768	17.978	141.42

An example equation for a negative binomial regression model in the Table above is as follows:

$$y = e^{(-16.1+0.157*PET\_1s-0.00193*Min\_SD+2.73*TCV)} \tag{6.13}$$

6.2 SUMMARY

For establishing the predictive power of PET, generalized linear models were developed using PET and other intersection characteristics (traffic volume, geometric characteristics etc.). The GLM analysis has shown that crash data has overdispersion and hence Negative Binomial model structure fits the data better than a Poisson model. More importantly, it showed that though a model combining PET and traffic volume characteristic (AADT or conflicting volume) has better predictive power than the ones containing only PET, the increase in complexity due to the inclusion of other intersection characteristics such as minor lane width improves the model fit only marginally if any, and does not justify such a complex model. Therefore, it can be concluded from chapter 5 and chapter 6 that PET has both predictive and diagnostic capability. PET can independently perform diagnosis of intersections by categorizing it as belonging to a safety category. When it comes to the ability to predict crashes, though the model combining exposure (AADT or conflicting volume) and PET has the lowest residual deviance, based on the principle of parsimony, it can be concluded that PET at 1 second threshold alone has significant predictive power. Inclusion of other intersection characteristics only marginally improves the model fit because PET in itself captures the effect of intersection characteristics on driver behavior and safety. Unsafe geometric conditions often make drivers accept risky gaps or react slowly to critical events, leading to serious conflicts. This phenomenon is captured by PET and hence PET exhibits good predictive power. This result agrees with the observations made in Kim et al., 2007 where they modeled crash types using hierarchical multilevel modeling approach: drivers' characteristics are nested within crashes, crash characteristics are nested within site characteristics, and site characteristics are nested within regional characteristics.

CHAPTER 7: INTERSECTION SAFETY SURVEY

7.1 INTRODUCTION

With varying levels of success expert knowledge based systems have been used in various studies related to road safety.).The human expert knowledge has been used either to supplement and enhance quantitative safety prediction systems for identification of hazardous locations and potential countermeasures (Herland et. al., (2000), (Tarko and DeSalle (2002)), or to test and validate already build hazardous-location identification systems (also called as expert systems in previous studies although human expert knowledge was used to test and validate (Spring et. al., (1991)). The former implementation attempts to capitalize on the likelihood that safety predictions may be improved by incorporating both quantitative and qualitative factors, while the later believes that qualitative evaluation by human experts is as important as evaluations based on quantitative measures. For example, Spring et. al., (1991) developed a prototype expert system called Hazardous Location Analyst (HLA) to analyze high-accident locations. This system was tested, verified, and validated against human expert inferences. Their paper suggests that a qualitative evaluation is as appropriate and important as quantitative performance measures for evaluating safety expert systems. In that effort it was seen that expert knowledge carries valuable information for applying safety knowledge to constraints of specific locations.

Herland et. al., (2000) developed a knowledge-based local traffic safety support system (KLOTS) that suggests countermeasures to traffic engineers to address safety problems. The system uses a set of rules to analyze the safety problem and then queries its knowledge-base to determine countermeasures, including providing pros and cons of the suggested measures. The knowledge-base is developed by conducting case interviews of experts on real-life traffic safety problems. An interesting observation in the KLOTS study is that there was a consensus found among the expert opinions in practically every area which improves the confidence in this approach.

Tarko and DeSalle (2002) used motorist feedback to identify hazardous intersections and supplement the use of crash data for such identification. The responses from the motorists were compared with the actual crash data. The study found that motorists have a good perception of hazardousness of locations. On an assumption that a location is considered hazardous if it had greater than 15 crashes in 3 years, it was found that respondents identified 45% of such locations. It was also found that 55 percent of locations reported by motorists were found hazardous. Consensus among motorists indicated better identification of hazardous location as this rate improved to 86 percent for locations reported at least twice, and to 96 percent for locations reported at least three times. Though this study was encouraging as a method to identify hazardous locations, its applicability is dependent on the feedback obtained. As the responses are mostly based on personal non-crash experience a motorist who had no or limited hazardous experiences at an intersection might consider the intersection safe. Thus identifying hazardous location likely becomes be dependent on a sufficient number of responses.

Expert knowledge has also been exploited in various other transportation areas. Cafiso et al. (2012) investigated the use of expert opinion in evaluating safety of bus transport in Italy. They applied a method called the Delphi technique. This technique was devised in 1950s as a means of handling opinions rather than objective facts. The method has two stages of using expert opinion to draw conclusions: combine and refine. In the refine stage, experts are revealed each other's opinions. They then reevaluate the given scenario knowing their expert colleagues' opinions. While a lack of consensus in terms of priority ranking of risks was found even after the refine stage, safety issues related to bus operations were highlighted along with the effectiveness of new technologies to improve safety. Another study was conducted by Rogerson and Lambert (2012) that exploited expert opinion to rank factors of safety which they called "factor hierarchies" related to airport runway incursions. They described a method for evaluating a set of locations based on varied factors of safety and levels of risks across them. Merat et al. (2011) compared the merits of expert observation, drivers' own assessment, and objective driving data to determine the effects of IVIS (in-vehicle information systems) on driving performance. All three methods were able to identify that using IVIS is a demanding task and will affect driving performance. However, the expert observation captures some of the unsafe longitudinal maneuvers of drivers resulting from the IVIS that was not be assessed by drivers as unsafe. This paper demonstrated the utility of using expert opinion.

During phase (ii) of the current research, a visual inspection of the intersections of Henry Clower Blvd/Oak Rd at GA 10 and Grayson Pkwy at GA 10 by an expert showed that the former was judged to be safer of the two intersections. This judgment is corroborated by the crash data. While anecdotal, this nevertheless suggests that there may be an additional factor(s) not captured in the current quantitative analysis that contribute to the probability of a critical event converting into a crash (or probability of a PET value becoming zero). As seen in the preceding discussion expert judgment may reflect this probability difference at the two intersections, allowing the expert to conclude that one intersection is safer than the other. That is, an expert in transportation safety mentally synthesizes the visual information available at the sites to determine the relative safety levels. This chapter will focus on exploring how well expert based evaluations compare to crash data for the 18 intersection presented in the previous chapter.

7.2 OBJECTIVE

As stated the objective of this effort is to quantify how well self-selected safety experts may identify the relative safety performance of an intersection. The specific evaluation will be based on the propensity of left-turn vehicle crashes with opposing through vehicles, when given basic visual, traffic, and geometric information about the intersection. To accomplish this, a survey was developed in which the participants are asked to assess four intersections. For each intersection, they are provided basic information, such as AADTs, estimated approach grades, images, videos, and a Google street view to assist them in their assessment. First they are asked to categorize each of these four intersections as having high, medium, or low propensity of left-turn vehicle crashes with through vehicles. As part of this categorization they are asked to select the factors that they believe are important in their assessment. Once they complete this assessment for each of the four intersections, they are then asked to provide an ordinal ranking for the four intersections based on the propensity of left-turn vehicle crashes with through vehicles. Survey participants are allowed to take this survey multiple times with a different set of intersections (maximum of five times) if they wish to do so. In addition, participants are sought with expert safety knowledge through the use of professional national safety committee memberships and other recommended means. However, knowledge is likely highly varied among participants and contacted individuals must self-assess if they have sufficient expertise to participate.

All pictures of survey pages used in this chapter are from the survey website (http://transposurvey.ce.gatech.edu/webSurvey/login/loginUDL.php?surveyCode=UDL5694LPN759VVRG) (Photo credit: Lakshmi (2013))

7.3 SURVEY STEPS

The first page of the survey shows the user images of four intersections randomly selected from the 18 candidate intersections analyzed in the previous chapter (Figure 7.1). Each participant receives a new random draw of four intersections out of the 18 candidates.

Figure 7.1: Example intersections assigned to the respondent

The user is first asked to provide individual assessment of each of these intersections, rating each as having a high, medium, or low propensity of left-turn vehicle crashes with through vehicles. To complete the assessment the user clicks on an intersection's name or image for additional information and the assessment form. As shown in Figure 7.2, an individual intersection's page contains some intersection characteristics such as AADT, grade etc.

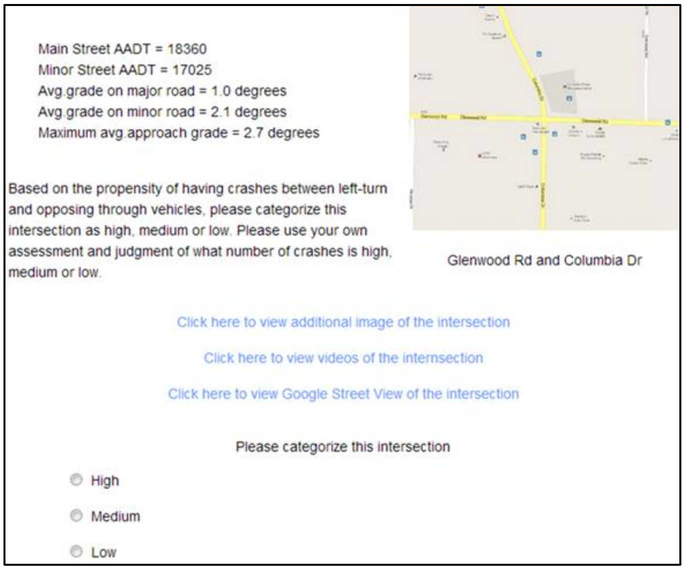


Figure 7.2: Example intersection page

Figure 7.3: Additional images of the example intersection

The survey also provides links to additional images of the intersection (Figure 7.3), recorded video clips, (Figure 7.4), and a link to Google street view of the intersection (Figure 7.5).

Figure 7.4: Videos recorded at the example intersection

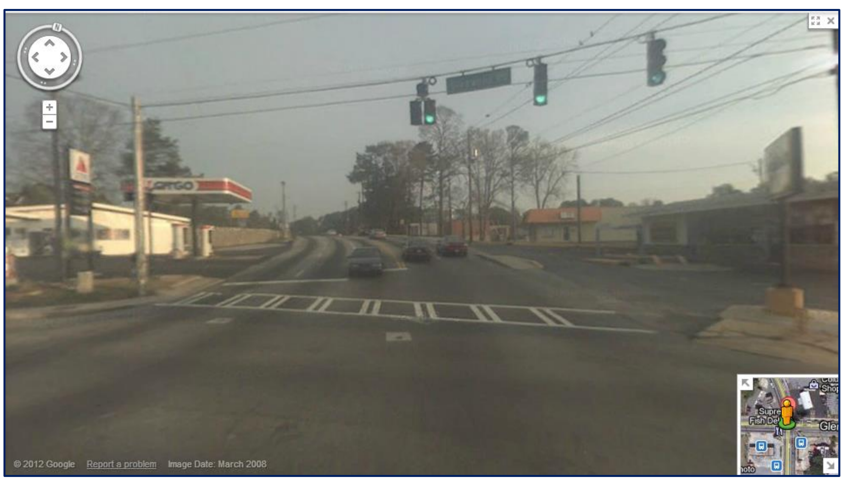
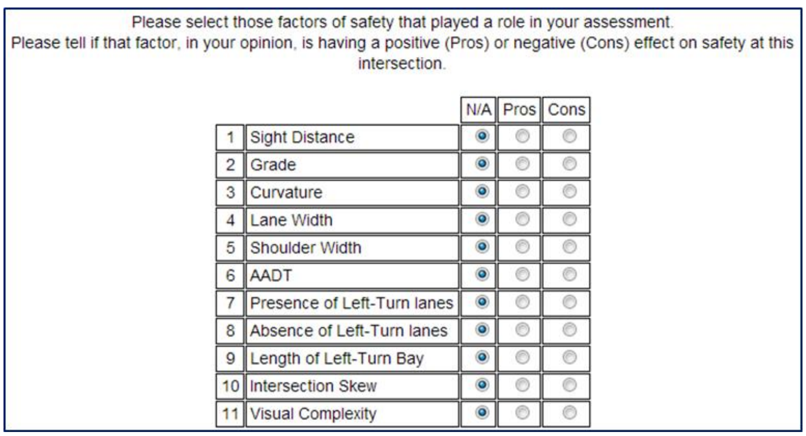


Figure 7.5: Google street view of the example intersection

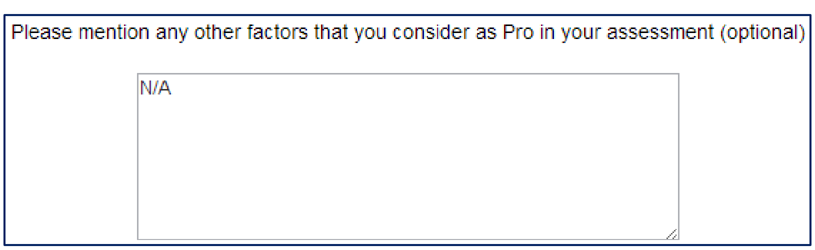
Once the user chooses the hazard category of the intersection, the next step is to determine what factors guided the user's evaluation. The respondent is provided with a list of potential factors of safety pertinent to left-turn vehicles (Figure 7.6). The default choice is N/A which means that according to the evaluator, the factor does not have any effect on the propensity of left-turn opposing crashes. If the user believes that a factor (for e.g. sight distance) made the intersection more hazardous they would select “Cons” for that factor. If the user feels that a factor (for e.g. presence of left turn lane) contributed to improving safety of left-turn vehicles they would select “Pros” for that factor. The next two steps ask the users if they had any other factors in addition to those listed that guided their evaluation (Figure 7.7 and Figure 7.8).



Please select those factors of safety that played a role in your assessment.
Please tell if that factor, in your opinion, is having a positive (Pros) or negative (Cons) effect on safety at this intersection.

		N/A	Pros	Cons
1	Sight Distance	◉	○	○
2	Grade	◉	○	○
3	Curvature	◉	○	○
4	Lane Width	◉	○	○
5	Shoulder Width	◉	○	○
6	AADT	◉	○	○
7	Presence of Left-Turn lanes	◉	○	○
8	Absence of Left-Turn lanes	◉	○	○
9	Length of Left-Turn Bay	◉	○	○
10	Intersection Skew	◉	○	○
11	Visual Complexity	◉	○	○

Figure 7.6: Factors of safety



Please mention any other factors that you consider as Pro in your assessment (optional)

N/A

Figure 7.7: Additional factors that are considered to be "Pro"

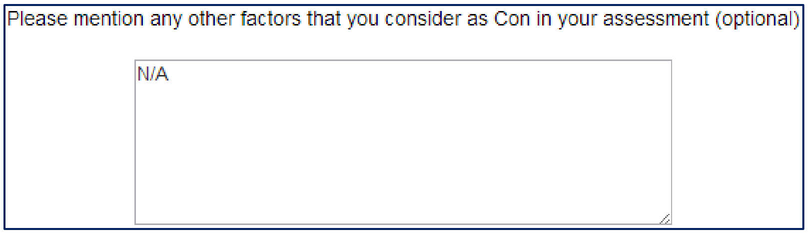


Figure 7.8: Additional factors that are considered to be "Con"

Performing the above steps completes the evaluation of that intersection. The survey then returns the user to the page showing the four intersections with intersection whose evaluation is completed “faded out” in relation to others (Figure 7.9). The user selects another intersection and repeats the process of evaluation.

Figure 7.9: Page showing completion of evaluation of two intersections

Once the user completes the individual evaluations of all four intersections they are asked to evaluate the intersections relative to each other (on a scale of 1 to 4 with 1 being the most hazardous and 4 being the least) (Figure 7.10). Each ordinal rank may only be used once.

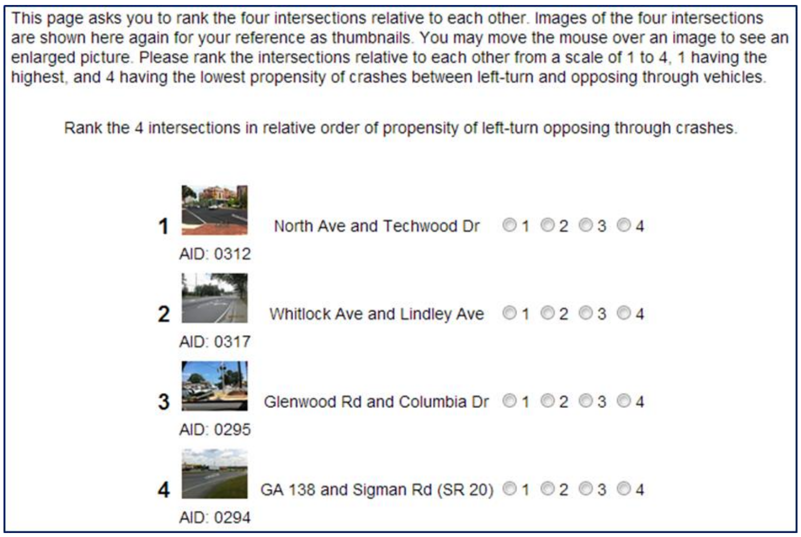


Figure 7.10: Relative evaluation of intersections

The survey also enables each user to repeat the survey four more times with a different set of four intersections assigned to them each time, if they wish to repeat the evaluation exercise.

7.4 SURVEY RESULTS

The survey was responded by 25 experts. Since the survey is anonymous, it is not possible to know if any of the experts completed the survey more than once. However, even if an expert completed more than one survey, random intersection assignment would make sure the intersections would not be repeated. Not all respondents completed all four intersections presented to them. Out of 100 possible intersection categorizations (4 intersections in each survey), only 94 intersections were categorized. Moreover, not everyone who completed the 4 intersections responded to the relative ranking question. Out of the 25 respondents, only 15 responses consisted of relative ranking of the four intersections presented to them. The following results need to be looked in this context.

The results from this survey are expected to inform three questions:

- How do the expert-based and crash based categorizations compare with each other?
- How does the experts' relative ranking of intersections (i.e. 1 through 4) compare with a crash based ranking?
- What are the factors that the experts identified as most important in the evaluation process?

7.4.1 Individual evaluation of intersections

The respondents are first asked to evaluate and categorize each intersection presented to them as belonging to either high, medium, or low crash category. Table 7.1 shows the results of individual evaluation of intersections by experts. This part of analysis helps us to evaluate the agreement between categorization based on crash data versus expert's evaluation. While the crash count is an objective measure, the categorization based on crash count is considered subjective as the category boundaries were selected by the analyst. Other groupings are clearly possible. For example, this study included no intersections with zero crashers. If zero crash intersection were included a categorization could have been created that defined intersections with zero incidents as low, defined all intersections with 1 to 50 as medium, and intersection with greater that fifty crashers as high. Another alternative could change thresholds, such as using 20 crashes as opposed to 50 to identify the high crash category.

Table 7.1: Individual evaluation of intersections by experts

Crash Count	Category (based on crash count)	Intersection Name	Categorization by experts							
118	H	N Druid Hills Rd and Lavista Rd	H	H	M	H	M	H	H	
90	H	GA 138 and Sigman Rd (SR 20)	H	H	M	M	M			
78	H	Roswell Rd and W Wieuca Rd	M	H	H	M				
73	H	Lawrenceville Hwy and Lawrenceville Suwanee Rd	M	L	L	M				
64	H	GA 20 and Willow Lane	M	M	M	L	M	M	H	
53	H	Grayson Hwy and Scenic Hwy	L	L	M	M	M	H	H	
48	M	N Druid Hills Rd and Lawrenceville Hwy	M	M	M	M	M	M	M	M
29	M	GA 10 and Grayson Pkwy	M	M	H					
27	M	Ponce de leon ave and Moreland ave	M	H	H	M	M			
23	M	Scott Blvd and Clairemont Rd	M	H	H	M	H	M	M	
15	M	Memorial Rd and Covington Hwy	L	M	L					
15	M	Glenwood Rd and Columbia Dr	M	M	H	M	L			
9	L	GA 10 and Oak Rd	M	M	L	L	L	M	L	M
6	L	Sugarloaf Pkwy and Buford Hwy	M	M	L	M	L	L	M	L
5	L	Cobb Pkwy and Gresham Rd	M	M	H	H				
2	L	MLK Jr Blvd and Brownlee Rd	L	M	L	M				
2	L	North Ave and Techwood Dr	M	M						
2	L	Whitlock Ave and Lindley Ave	M	M	L					

As an initial observation it can be seen that the number of responses for each intersection are not the same. The reason is that the intersections are randomly chosen for presentation to the responder for evaluation, with no guarantee of a balanced number of intersections samples over the experiment. A review at the Table 7.1 shows that an intersection belonging to the crash based “High” category was presented 34 times, crash based “Medium” category was presented 31 times, and crash based “Low” category was presented 29 times. There are 55 responses evaluating an intersection as “Medium”, 21 responses evaluating an intersection as “High”, and only 18 responses evaluating an intersection as “Low”. The number of “Medium” categorizations is proportionally much highly under the expert categorization than the crash categorization. Table 7.2 shows that the number of medium categorizations by the experts is almost equal (≈ 50 %) for intersections belonging to both high and low crash categories. The probability of experts categorizing a crash based high category intersection as low or crash based low category intersection as high is minimal (14.8% and 6.8 % respectively). This demonstrates, for the intersections given, a strong tendency of the experts to the medium classification. This may reflect a tendency of experts to default to a medium ranking. However, it may also reflect a bias in the crash based categorization implying the set threshold did not adequately categorize intersections.

Table 7.2: Aggregate numbers depicting categorization by experts

	Expert Categorization (% of crash categorization)		
Crash Categorization (Number of Responses and %)	H	M	L
High (34)(36.1%)	12 (35.2%)	17 (50%)	5 (14.8%)
Medium (31)(32.9%)	7 (22.5%)	21 (67.7%)	3 (9.6%)
Low (29)(30.8%)	2 (6.8%)	17 (58.6%)	10 (34.4%)

In order to obtain a deeper understanding of the evaluations of intersections by experts, a basic statistical analysis has been presented in Figure 7.13. Each evaluation of an intersection as “High” has been given a weight of 3, each “Medium” has been given a weight of 2, and each “Low” a weight of 1. Finally, the mean and standard deviation of these weights for each intersection has been calculated and presented.

Table 7.3: Variance in evaluation of intersections by experts

Crash Count	Category (based on crash count)	Weight of Selected Category								Mean	SD
118	H	3	3	2	3	2	3	3		2.71	0.49
78	H	2	3	3	2					2.50	0.58
73	H	2	1	1	2					1.50	0.58
90	H	3	3	2	2	2				2.40	0.55
64	H	2	2	2	1	2	2	3		2.00	0.58
53	H	1	1	2	2	2	3	3		2.00	0.82
48	M	2	2	2	2	2	2	2	2	2.00	0.00
29	M	2	2	3						2.33	0.58
27	M	2	3	3						2.67	0.58
23	M	2	3	3	2	3	2	2		2.43	0.53
15	M	1	2	1						1.33	0.58
9	M	2	2	1	1	1	2	1	2	1.50	0.53
5	L	2	2	3	3					2.50	0.58
15	L	2	2	3	2	1				2.00	0.71
2	L	1	2	1	2					1.50	0.58
6	L	2	2	1	2	1	1	2	1	1.50	0.53
2	L	2	2							2.00	0.00
2	L	2	2	1						1.67	0.58

The above Figure shows that the mean weight for the intersection having the highest number of crashes is 2.71 (expected a mean of 3) which demonstrates good agreement between the expert and crash categorizations. The second and third intersections also have high average values at 2.4 and 2.5, respectively. However, the next intersection, having 73 crashes, shows high level of disagreement between crashes and expert based categorization. According to crash numbers the intersection is in the high category, but none of the four respondents assigned a high categorization. Similarly, for low crash intersection all other intersections have a mean value of 2 or less, except for one intersection that has a mean rating of 2.5. This again shows some degree of agreement between categorization by experts and crashes. The above analysis also shows that the experts are more inclined to categorize an intersection as “Medium” unless it is clearly evident (as in the case of Lavista Rd. and N. Druid Hills Rd.). Assuming the same weights for the intersection categorizations as well (3 for H, 2 for M, and 1 for L), a Wilcoxon Matched Pairs Test was conducted to test the agreement between the categorization. The p-value for the test was found to be 0.88, which means that there is no evidence to reject the null hypothesis that the medians of the two datasets differ (in other words agreement between crash categorization and expert categorization). However, a caveat is that for small samples, the Wilcoxon Matced Pairs Test has little power to detect small differences. Scatter plot (Figure 7.11) between mean rating by experts and crash frequency shows a weak relationship with an R^2 of 0.2.

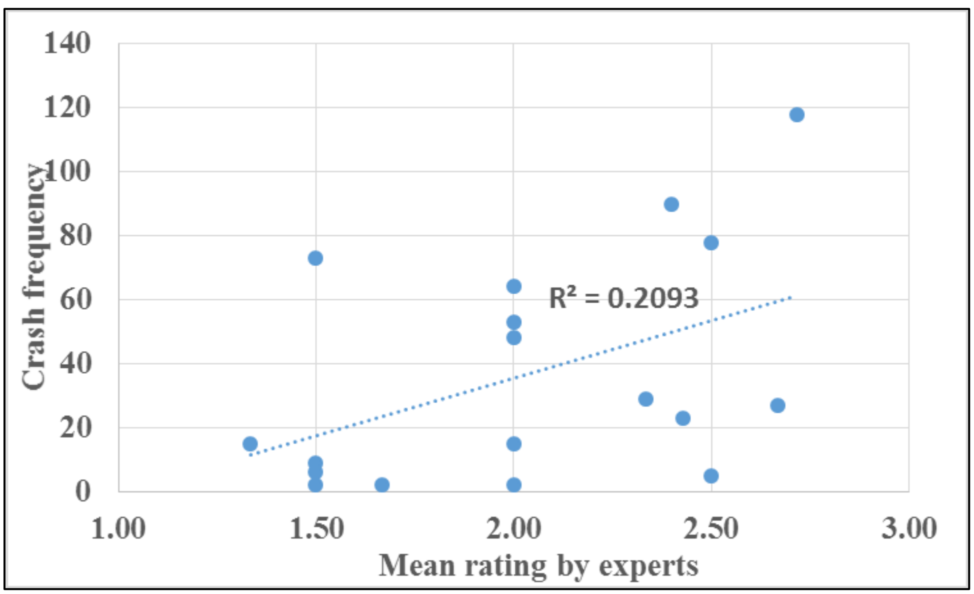


Figure 7.11: Relationship between expert categorization and crash categorization

Wilcoxon Matched Pairs Test was also conducted on PET parameters to test if their rank ordering agrees with that of experts. Table 7.4 shows the corresponding p-values for this test. The test shows that only PET_1s data has agreement with expert classification but

PET values at all other threshold do not. The possible reason for this is that experts have the ability to filter out high crash intersections better than medium-low crash intersections, and it is already established in Chapter 5 that PET_1s also has the same ability. Figure 7.12 shows the relationship between mean rating by experts and PET measure. It can be seen that the evaluation by the experts agrees and PET measure have a low linear correlation.

Table 7.4: Results of rank test between expert categorization and PET frequency

Parameter	p-value
PET_1s	0.25
PET_1.5s	<0.001
PET_2s	<0.001
PET_2.5s	<0.001
PET_3s	<0.001

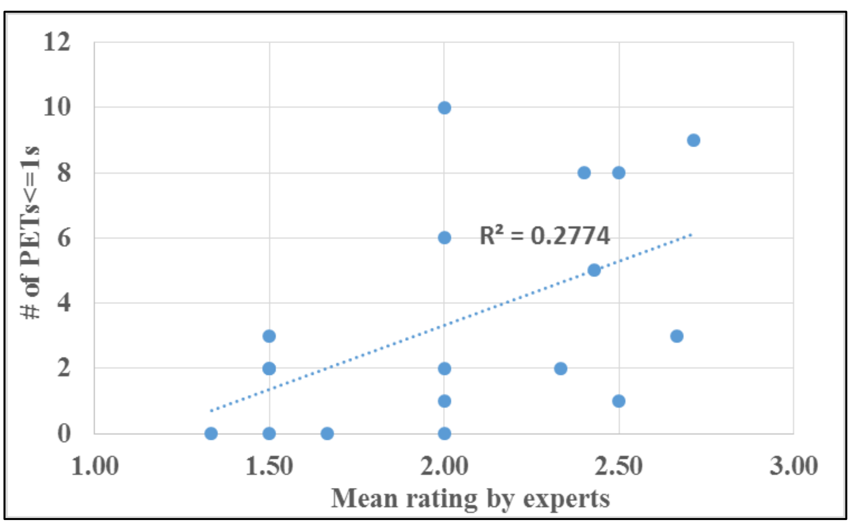


Figure 7.12: Relationship between expert categorization and PET frequency

In the survey we asked the respondents to consider only the left-turn vs. opposing through crashes. Though images and videos of these intersections were provided in the survey, it may be difficult for the experts to differentiate between overall hazardousness of the intersections and hazardousness specific to crashes between left turns and opposing through vehicles. Therefore the analysis shown in Table 7.3 is repeated with total crashes at the 18 study intersections. The results are shown in Table 7.5. A Wilcoxon Matched Pairs Test however, gave a p-value of <0.0001 which means there is sufficient evidence to reject the null hypothesis that there is agreement between the ranks of total crashes and categorization by the experts. In fact, this is a good finding because it shows that experts did consider the hazardousness specific to left turn vs. opposing through crashes.

Table 7.5: Variance in evaluation of intersections by experts with respect to total crashes at the study intersections

Total Crashes	Intersection Name	Mean categorization by experts	SD of categorization by experts
429	GA 138 and Sigman Rd (SR 20)	2.40	0.55
392	GA 20 and Willow Lane	2.00	0.58
256	Grayson Hwy and Scenic Hwy	2.00	0.82
243	N Druid Hills Rd and Lavista Rd	2.71	0.49
243	N Druid Hills Rd and Lawrenceville Hwy	2.00	0.00
207	Lawrenceville Hwy and Lawrenceville Suwanee Rd	1.50	0.58
190	Ponce de leon ave and Moreland ave	2.67	0.58
173	Glenwood Rd and Columbia Dr	2.00	0.71
167	Roswell Rd and W Wieuca Rd	2.50	0.58
158	GA 10 and Grayson Pkwy	2.33	0.58
140	Memorial Rd and Covington Hwy	1.33	0.58
96	GA 10 and HenryClowerBlvd - Oak Rd	1.50	0.53
94	Sugarloaf Pkwy and Buford Hwy	1.50	0.53
79	Scott Blvd and Clairemont Rd	2.43	0.53
51	Whitlock Ave and Lindley Ave	1.67	0.58
46	Cobb Pkwy and Gresham Rd	2.50	0.58
29	North Ave and Techwood Dr	2.00	0.00
26	MLK Jr Blvd and Brownlee Rd	1.50	0.58

7.4.2 Relative evaluation of intersections

The previous subsection 7.4.1 explored the agreement between crash based categorization and expert based evaluation based on the high, medium, and low categorization of each individual intersection. As the data showed, rank based test has shown agreement but low linear correlation value showed only partial agreement. However, an absolute categorization of an intersection as belonging to a crash category may not be necessary. It may suffice to use expert opinion in ranking given intersections relative crash propensity. This subsection analyzes this aspect of expert evaluation.

Details of relative rankings for each group of intersections (based on crash numbers and expert ranking) are provided in the Appendix F. One way of establishing the similarity or the lack thereof, of the relative rankings between those provided by the experts and that indicated by the crash numbers is by computing a correlation coefficient between the ranks. Figure 7.13 shows the summary results of the computed correlation coefficients for each group of intersections provided to each expert.

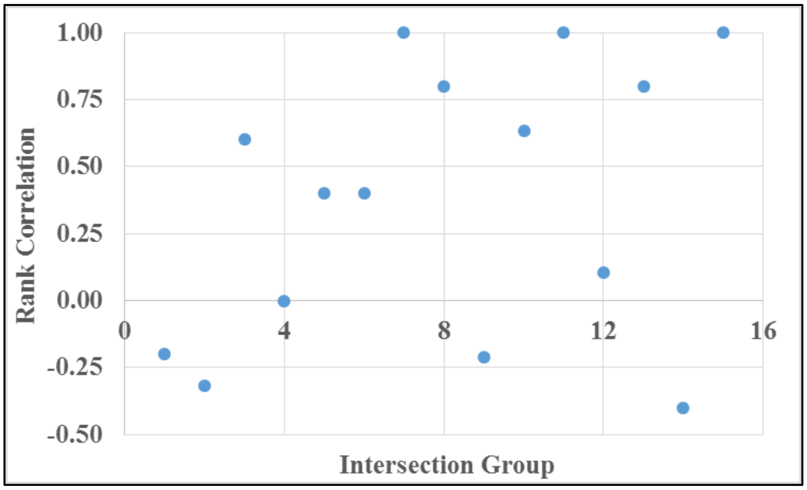


Figure 7.13: Relative ranking among intersections

The number of intersection groups considered in this analysis is lesser than the total number of respondents considered in subsection 7.4.1 because 10 of the 25 respondents did not provide the relative rankings. This plot shows that a majority of the correlation coefficients are positive, 5 of them are more than 0.75, out of which 3 show perfect correlation, which alludes to some agreement between the relative rankings of intersections provided by the experts and indicated by crash data. A deeper look at the cases where there was significant disagreement between relative rankings provided by experts and crash data, two specific cases emerge.

- In an example case (Table 7.6) that had a negative rank correlations, it can be seen that though there was only one difference in the relative ranking, the coefficient was still negative. In this example we can see that though 3 of the intersections' relative rankings are same for both experts and crash data, ((1,2,3), there is a negative correlation as the expert ranked the intersection with the fewest crashes as the least safe. The reason for this is that rank values are not completely independent values (all the ranks are bounded, in this case between 1 through 4). The interpretation of rank correlation values should be made in this context in comparison to correlation coefficient (R^2).

Table 7.6: Example case of negative correlation between ranking by experts and crashes

Intersection	Expert Ranking	Crash Ranking
Grayson Hwy and Scenic Hwy	4	1
Ponce De Leon Ave and Moreland Ave	1	2
Scott Blvd and Clairemont Ave	2	3
Sugarloaf Pkwy and Buford Hwy	3	4

A Wilcoxon Matched Pairs test was also done on these pairs of relative rankings. Since, a sample of 4 in each group is too small for the test, all the pairs are combined in one single test. The p-value for this test is 0.99 which shows a strong agreement between the relative rankings between experts and crash numbers.

- When intersections are presented to the respondents that belong to the same crash based category (Table 7.7) this often results in lower rank correlation values. In cases of intersections that belong to different crash categories, correlation of rankings between those provided by experts and crash data is better. This means that even experts find it difficult to relative rank intersections that are similar in their risk categorization for opposing left-turn crashes. This is not necessarily a bad finding because ability to relatively rank within a category is not as important (especially for low and medium crash intersections) as classification of an intersection. So, the rank correlation results might vary based on the combination of 4 intersections presented to the experts.

Table 7.7: Example case of relative rankings of intersections

Intersection	Crash Categorization	Expert Ranking	Crash Ranking
Glenwood Rd and Columbia Dr	L	1	2.5
N. Druid Hills Rd and Lawrenceville Hwy	M	3	1
North Ave and Techwood Dr	L	2	4
Sugarloaf Pkwy and Buford Hwy	L	4	2.5

7.4.3 Factors of safety

The next section explores the factors of safety that the experts identified as important in their assessment. The results from this section of the survey are presented in Table 7.8. The responses of experts with respect to each factor of safety have been classified based on the safety category of the intersection. Each factor of safety has three possible selections – if the factor is a pro, a con, or NA (if the experts think that the factor is neither a pro nor a con).

Table 7.8: Selection of "factors of safety" by experts

Intersection Category (by expert)	Factor Class	Sight Distance	Grade	Curvature	Lane Width	Shoulder Width	AADT	Presence of Left-turn Lanes	Absence of Left-turn lanes	Length of Left-turn bay	Intersection Skew	Visual Complexity	Row Total (%)
H	NA	12	6	7	10	2	20	6	15	10	6	2	96 (41.5%)
	Pro	1	6	9	0	19	1	3	0	2	2	5	48 (20.7%)
	Con	8	9	5	11	0	0	12	6	9	13	14	87 (37.6%)
M	NA	32	37	39	28	29	47	39	43	34	25	32	385 (63.6%)
	Pro	13	10	11	9	20	5	11	8	8	11	6	112 (18.5%)
	Con	10	8	5	18	6	3	5	4	13	19	17	108 (17.8%)
L	NA	8	10	7	10	5	16	7	9	13	7	6	98 (49.4%)
	Pro	9	7	11	6	11	2	8	6	3	10	5	78 (39.3%)
	Con	1	1	0	2	2	0	3	3	2	1	7	22 (11.1%)
Column Total	NA	52 (53.4%)	53 (56.4%)	53 (56.4%)	48 (51%)	36 (38.3%)	83 (88.3%)	52 (55.3%)	67 (71.3%)	57 (60.6%)	38 (40.4%)	40 (42.6%)	
	Pro	23 (24.4%)	23 (24.4%)	31 (32.9%)	15 (16%)	50 (53.2%)	8 (8.5%)	22 (23.4%)	14 (14.9%)	13 (13.9%)	23 (24.5%)	16 (17%)	
	Con	19 (20.2%)	18 (19.2%)	10 (10.7%)	31 (33%)	8 (8.5%)	3 (3.2%)	20 (21.3%)	13 (13.8%)	24 (25.5%)	33 (35.1%)	38 (40.4%)	

First of all, though NA forms the majority of responses, the percentage of "cons" for high crash intersections and percentage of "pros" for low crash intersections are higher than the other, which is according to expectations. "Visual Complexity" comes out as the most significant "con" among all three categories of intersections. "Intersection Skew" and "Lane Width" are also significant factors that act as "cons" for high and medium crash intersections, while they become "pros" for low crash intersections (which means wider lanes and absence of intersection skew). In addition "length of left-turn bay", "grade" and "curvature" are the other significant "cons" for high crash intersections. For the low crash category intersections, there are fewer selection of factors as "cons", which is expected. The factors "sight distance", "curvature", "shoulder width", and "presence of left-turn lanes" are considered to be substantial "pros" for low crash intersections..

For the factors that are considered as a "Pro", for all the three crash categories, the factor "Shoulder Width" was chosen very frequently (19, 20, and 11 times respectively, aggregating to more than 50% of responses) making it a significant factor in the opinion of the experts. “Grade”, and “curvature” are two other important “pro” factors selected by the experts. It can be interpreted as that the lack of these factors is a “pro” for safety of intersections. For low crash intersection category, “sight distance” is also a significant “pro” factor. However, it is interesting to note that for all the three categories, the experts have unanimously agreed that “AADT” is not a significant factor in their assessment of safety between left-turning and opposing through vehicles, given that literature shows that AADT is an important factor of safety. A potential reason might be that the magnitude of AADT at an intersection cannot be directly visualized as well as some other factors of safety that are static in nature such as lane width, intersection skew among others.

Overall, the most significant factors that the experts felt influenced their evaluation of intersections can be determined by counting the total number of selections (either a “Pro” or a “Con”) for each. The greater the number of selections, the greater the perceived significance. Thus, the most significant factor appears to be “Shoulder Width” with 58 selections, “Intersection Skew” with a total of 56 selections, followed by “Visual Complexity” with 54 selections. However, given that traffic volume as a measure of exposure is an important factor of safety, it is surprising to see that the factor “AADT” comes out as the least significant factor effecting expert evaluation, with only 11 selections.

7.4.4 Additional Factors

The next section of analysis deals with additional factors recognized by experts that influenced their evaluation of intersections. All these comments are presented concisely in the form of a factor of safety as shown in Table 7.9. Though the experts might not have used the same words mentioned in the Figure below, they have been categorized into their closely matching factors. It can be seen that some of these factors are already presented in the survey as selections under “pros” or “cons”. However, some experts still chose to mention them (probably to signify their importance).

It can be seen that the highest number of mentions have been for the category of “Speed”. Speed limits corresponding to every intersection were provided as metadata in the survey. It is possible that the experts based their judgment on the combination of speed limit, and actual movement of vehicles. The mention of this factor either has been in the form of magnitude of speed of oncoming vehicles, or the problem of misjudging speed of oncoming vehicles. This implies that speed of through vehicles does count as an important factor in safety of left-turn vehicles. Interestingly, “Protected left-turn phasing” is the second highest mentioned factor where either the experts felt that this would increase safety of left-turn vehicles or that they would take lesser risks during permissive phase especially during peak hours. The third highest mentions fall into the category of “Obstructing sight/ Good sight distance”. Even the next highest mentions of “Congestion/Complexity” and “Grade” are related to the already mentioned factors such as sight distance or judging the speed of oncoming vehicles.

Table 7.9: Additional "Pros" and "Cons"

Factor	Number of Mentions
Speed	15
Protected left-turn signal phasing	10
Obstructing sight/ Good sight distance	9
Congestion/Complexity	8
Grade	8
Open intersection	7
Volume on minor road	6
Number of lanes	5
Channelization	4
Access control	4
Buffer lane/ Buffer space	4
Intersection on slope	3
Pedestrians	3
Pavement markings	3
Skew	2
Intersection of state routes	1
Intersection size	1
Rural intersection	1
Trees	1

7.5 SUMMARY

This chapter explores the qualitative aspect of safety evaluation by attempting to use expert knowledge in assessing the propensity of left-turn vehicle crashes with opposing through vehicles at intersections. The results of the utilized survey show that the evaluations done by the experts partially agree with those shown by crash data. At an aggregate level, it was observed that the experts are more prone to classifying an intersection as when compared to the crash based categorization. The proportion of "medium" categorizations assigned by experts for intersections that belong to low and high crash categorizations is almost equal. Even though Wilcoxon Matched Pairs test showed evidence of agreement between expert categorization, crash categorization, and

PET_1s, it was also seen that the linear correlation between expert categorization and PET measures is slightly higher than that found between expert categorization and crash categorization. This shows a promise in combining quantitative (crashes, surrogate) and qualitative measures for safety evaluation in future efforts.

The next observation is that the relative ranking done by the experts also agrees with relative ranking based on crash numbers. The relative ranking is much better when the evaluated intersections have varied safety in comparison to a group of intersections that belong to the same safety category. Also, in practice, relative ranking of intersections is sought more than an absolute categorization of each intersection. So, there is an applicability of expert evaluation in case of relative rankings in terms of crash propensity.

The next section investigated what factors of safety the experts stated as important in the safety evaluations. The most significant “con” factor appears to be “Intersection Skew” with a total of 56 selections, followed by “Visual Complexity” with 54 selections. “Shoulder Width” with more than 50 “pro” selections is the most significant “pro” factor as per experts. However, given that traffic volume as a measure of exposure is an important factor of safety, it is surprising to see that the factor “AADT” was identified as the least significant factor effecting expert evaluation, with only 10 selections. Finally, it is found that “speed” is a factor that is mentioned the most number of times by experts. Future studies could take this factor into account in the modeling process. This exercise has shown that expert categorizations partially agree with crash data categorization and PET data. The possibility of combining qualitative and quantitative measures in developing safety models should be considered in future research. Moreover, this chapter had a limited number of experts (25) who have responded limiting the data available for analysis. The observations and conclusions might be different in case a higher number of experts respond in future.

CHAPTER 8: CONCLUSIONS AND RECOMMENDATIONS

8.1 RESEARCH SUMMARY

The current research is aimed at the evaluation of potential surrogate safety data collection methodologies and the effectiveness of several surrogate measures. A surrogate safety measure is defined as measurable or observable non-crash event that can either be converted or calibrated to crash frequency (Tarko et al., 2009),. As part of this evaluation the research also considered statistical sufficiency of data sample for each of these surrogates and guidance on their broader applicability. This effort focused on the crossing events between left-turning vehicles and opposing through vehicles leading to opposing left-turn crashes. The overall research has been divided into various sub research phases that have a logical connection.

The first phase of this research focused on three surrogate measures namely acceleration/deceleration values of through vehicles, intersection-entering speed of through vehicles, and post encroachment time (PET). These surrogates were collected and analyzed in the context of evaluating the effectiveness of a treatment aimed at improving the safety of interactions between left turning vehicles and opposing through vehicles. Conflicts were expected to be reflected by the observation of high acceleration/deceleration values or low PET values. PET in this context is defined as the time between the moment a left-turn vehicle leaves the area of encroachment and the moment when the first conflicting through vehicle enter the conflict area. By definition, a PET value of zero implies a crash. The speed with which through vehicles enter the intersection was also considered to be a potential measure that would capture the effect of the treatment.

A semi-automatic method using custom software was developed for extracting the surrogate safety data from videos. The methodology requires collecting speed data at fixed intervals along a longer stretch of approach road. This semi-automatic approach allows for the use of lower camera angles with larger perspective views. This approach limits equipment needs, and enables profile-based data collection even during dense traffic conditions, still a limitation of most of the currently developed automatic video detection based approaches. The methodology involves plotting the speed profiles and acceleration-deceleration profiles of vehicles. The methodology developed successfully extracts speed, acceleration/deceleration profiles of individual vehicles, and PET values from video data. However, this phase showed that the methodology for collecting vehicle profile data remains time consuming and labor intensive. It has also shown that PET has a high likelihood of providing a usable and cost-effective surrogate measure. It is relatively easy to measure as it requires collecting only two timestamps for each PET data point and a PET value of zero differentiates crash and non-crash events. Therefore, it was decided that the next phase of the research would focus on evaluating the effectiveness of PET as a surrogate measure for safety.

The second phase of research delved deeper into the effectiveness of PET as a surrogate measure of safety. This phase involved data collection at additional intersections, evaluating the properties and distribution of PET, and finally evaluating its effectiveness as a surrogate measure. The first focus of this phase of research was to collect PET data at four intersections having varied crash histories but similar operating characteristics to see if PET can capture the differences in crashes. The analysis of collected data allowed for the development of additional hypotheses on the effectiveness and application of PET as a surrogate measure for safety. These hypotheses were explored in phase 3. In phase 3 PET data was collected at additional intersections for a total 18 intersection locations. The analysis of data found that PET threshold, defined as the PET value at or below which PET acts as a surrogate safety measure, is important in determining the effectiveness of PET as a surrogate safety measure. It was found that the threshold may lie as low as 1 second to identify potential high crash intersections. The research then moved on to establishing the diagnostic and predictive power of PET with respect to finding crash propensity or frequency. Non-parametric rank analysis methods and generalized linear modeling techniques were used to model PET with other intersection and traffic characteristics. The results are listed in the "findings and conclusions" section below. The effectiveness of PET and its assistance to decision makers has also been demonstrated through an example that identified potential errors in crash data. Overall this research demonstrated that PET can be an effective surrogate for crashes between left-turning vehicles and opposing through vehicles but only at very low threshold values.

Additionally, an intersection safety survey was also conducted to explore the use of "expert" opinion to evaluate safety of intersections. This survey tested if a group of safety experts can successfully synthesize visual and quantitative information about a set of intersections and then evaluate the relative safety levels of those intersections by categorizing them into different safety categories.

8.2 FINDINGS, CONCLUSIONS, AND CONTRIBUTIONS

- The first contribution of this research is the development of a methodology to collect surrogate data, and demonstration of the advantages and challenges associated with developing such a methodology. For example, the methodology developed in the first phase of the research to collect acceleration/deceleration values allows for the use of low camera angles with large perspective views thereby limiting equipment needs, a limitation of most of the automatic video detection equipment based approaches. This part of research also demonstrated the difficulties and limitations of collecting profile based surrogate data such as speed profiles and acceleration-deceleration profiles with respect to noise in the raw data and sampling. It developed and validated algorithms to filter the raw data collected so that such smoothed data can be used for further analysis.

- The second and third phases of this research contribute towards obtaining an in-depth understanding of the effectiveness, limitations, and applicability of PET, focusing on permitted left turns at signalized intersections. The second phase of research showed that PET data collected during peak and non-peak hours might vary greatly and it is important to have a data collection period comprising of both categories. Moreover, the second phase of research also led to certain hypotheses with respect to applicability of PET that guided research for the third phase.

- The effectiveness of PET and its assistance to decision makers has also been demonstrated through an example that helped identify errors in crash data. Most transportation funding agencies rely on the crash data to rank intersections and to fund projects. The above analysis shows that PET can act as a tool to guide decision makers and increase their confidence in identifying intersections that require safety treatments and funding.

- One of the major contributions of this research is advancing an understanding of the importance of a threshold value for using surrogates. Extensive literature review was conducted and presented in this respect to support this theory. Analysis of data collected at the sample intersections for this phase of research showed that PET has the best correlation with crashes at a threshold as low as 1 second. Due to absence of any previous study to establish the threshold value for the applicability of PET as a surrogate, arbitrary threshold values in the order of 3 seconds have been reported in the literature. This is an important finding because it advances the knowledge on the applicability of PET as a surrogate, and signifying its importance for using other surrogate measures in general.

- Another contribution of this research is to demonstrate the use of PET both as a diagnostic and predictive tool. In order to evaluate the diagnostic power, a non-parametric method called Fischer's test was conducted to determine PET's categorical classification ability. The test conducted on the sample of intersections considered for this research showed that while a PET threshold of 1.5 second or 1 second was applicable in determining high crash locations, threshold of 3.0 seconds was valid in identifying low crash intersections. In fact, sensitivity analysis discussed in Chapter 5 showed that the choice of PET threshold would depend on the specific application to classify intersections, and the parameters' utility would depend on the subjective definition of "high" category. For example, PET threshold of 3 seconds can be used to also filter high-crash intersections if the threshold for classifying an intersection as "high" is at or below 20 opposing left-turn crashes. The 3 seconds threshold loses its classifying ability for categorization of intersections having higher opposing left-turn crash counts. Further comparison of PET at 1.5 second and 1 second threshold showed that though both these thresholds can group high-crash intersections, 1 second threshold has a further ability to rank intersections within a category which is why it showed the best rank correlation with crashes.

- Assuming that the "high" category consists of intersections having greater than 50 opposing left-turn crashes in 4 years, the non-parametric analysis also identified threshold values for PET parameters that can be used to identify high crash locations among the intersections considered in this study. A simple classification tree analysis showed that a threshold of 6 for the parameter PET_1s (intersections having lesser greater than or equal to 6 PETs below 1 second in the 5-hour period of data collection as done in this research) and a threshold of 1.92% for CDF1 seem to be doing a good job of grouping and identifying high crash locations among all. A threshold of 2 for the parameter PET_1s (intersections having lesser than 2 PETs below 1 second) also has shown potential to identify low crash intersections.

- Empirical distribution analysis has shown that PET data shows evidence of following GEV distribution at the tail. When the observed CDF values were computed based on the conditional probability, given PET thresholds ranging from 4 seconds to 1.5 second with 0.5 second interval, the plots showed that the GEV distribution fitted on PET data matches observed PET CDF values until 1.5 second and starts to deviate at 1 second. Importantly, the plots also showed that crashes (probability for PET=0) were overestimated by the GEV fit by a factor of 30 to 40 when the distribution is fit with a threshold of 2.5 second and above, and by a factor of 14 at 2 second threshold. This firstly shows that PET values at higher thresholds (2 seconds and above) are not acceptable predictors of crashes. Secondly, the process of crash occurrence is likely different from occurrence of PET values of 2 seconds and above most probably due to driver intervention at low PET thresholds of 1 second and below where drivers perform maneuvers to avoid crash. Therefore it is possible that higher PET thresholds overestimate crashes. However, due to insufficient data at 1 second and below, the analysis could not be extended to PET values 1.5 second and below. More data at 1 second and below thresholds is required to make any conclusions about utility of these lower thresholds for crash prediction.

- It was observed that traffic volume (in the form of conflicting volume) is a better measure to identify low crash intersections in comparison to high crash intersections. Even scatter plots between traffic volume and crashes has shown two trend lines signifying the difference in the relationship for low crash and high crash intersections. This analysis has also shown that minor road AADT was found to have better rank correlation with crashes than major road AADT, while for conflicting volume, major conflicting volume has better correlation with crashes.

- The predictive power of PET was estimated by developing models using PET, traffic characteristics and intersection characteristics to predict crashes. Various types of Generalized Linear Models were tested. It was found that the hypothesis of "a negative binomial model with a log link function was a good fit" could not be rejected. Moreover, a model containing both PET and traffic volume characteristic gave the simplest best-fit model, though inclusion of traffic volume only marginally improves the fit over a model containing only PET (CDF1s). However, further inclusion of other intersection characteristics such as minor lane width does not improve the fit. The underlying intuition of using PET as a surrogate measure for crashes is that it reflects the behavior of drivers at an intersection which is impacted by intersection specific characteristics (geometric characteristics, visual complexity etc.) that cannot be captured by just using traffic volume parameters or other geometric parameters. The hypothesis is that PET can provide supplemental information about the intersection. It logically follows that a model including a combination of PET, and traffic volume characteristics is a better fit to predict crash frequency in comparison to using these parameters individually. Moreover, as PET is expected to capture the impacts of other intersection characteristics, adding these characteristics in the model does not improve the fit significantly.

- Another contribution of this research is to test if a group of self-identified safety experts can successfully synthesize visual and quantitative information about a set of intersections and then evaluate the relative safety levels of those intersections by categorizing them into different safety categories. This effort also shows what intersection characteristics and factors guided the experts' evaluation and what effect ("pro" or "con") each of these factors has on the safety of left-turn vehicles. The research also compares the safety categorization performed by the expert panel and that reflected by the crash data. At an aggregate level, it was observed that the experts are more prone to classifying an intersection as medium unless that intersection has a very low likelihood of crashes or a relatively high likelihood of crashers. The proportion of "medium" categorizations assigned by experts for intersections that belong to low and high crash categorizations is almost equal. It was also seen that the correlation between expert categorization and PET measures is higher than that found between expert categorization and crash categorization. This shows a promise in combining quantitative (crashes, surrogate) and qualitative measures for safety evaluation.

- Wilcoxon Matched Pairs tests were conducted to test the agreement in ranks between classification done by the experts, intersection ranking based on crash frequency, and

PET measure. It was found that expert ranking has statistically significant agreement with the rankings provided by crash numbers and PET_1s parameter. The relative ranking done by the experts agrees with crash rankings better than individual intersection safety categorization. The relative ranking is also better when the evaluated intersections have varied safety in comparison to a group of intersections that belong to the same safety category. This exercise demonstrated the applicability of human experts' feedback as a potential independent source for safety evaluations.

- Finally, it was found that "speed" is a factor that is mentioned often by experts as important in safety evaluation of opposing left-turn crash propensity. Future studies could take this factor into modeling process.

Finally, it is expected that this research will help to determine needs for future research in this area.

8.3 FUTURE RESEARCH

- One of the limitations of the described data collection methodology used in this research is that the data could not be collected during night time, as identification of the correct frame when vehicles reach the detection line would be very difficult. This limitation exists for any methodology that exists as of today for extracting vehicle trajectories using video data. Thus, one of the future research topics should be to develop methodologies to collect data during night time. Moreover, the first phase of the research has exposed the difficulties in collecting surrogate data from videos and a reliable method for automation of such processes through computer vision techniques would also greatly increase the ability to broadly implement surrogate measures.

- The analysis of PET has shown the importance of establishing a threshold value for a surrogate to use the surrogate effectively to estimate crash propensity. The current research has shown that for PET, the threshold lies as low as 1 second (for recognizing high crash intersections) and that a threshold of 1.5 has a higher power to classify low crash intersections than a threshold value of 3 seconds. Similarly, it is important to establish such threshold values for various other surrogates used in the previous studies before using such measures as surrogates.

- Results from the Fischer's Exact Test are based on the sample intersections considered in this research. However, given the heavy skew in the distribution of crash numbers across intersections in Atlanta metro, data from more intersections is required to test the observations from this research for the population of intersections.

- One of the limitations of the data collected for this research is that there are very few PET data points below 1 second due to data collection limitations, both in technical equipment and limited sample size. The empirical distribution analysis has shown that PET values of 2 seconds and above follow a GEV distribution that over-estimates crash frequency. At a threshold of 1.5 seconds and below a different distribution is followed by PET data. However, due to unavailability of sufficient data at 1 second and below, such analysis could not be extended. Collection of more data below 1 second in future research will give a greater understanding of the PET distribution approaching a value of 0..

- Though it can be generally surmised that threshold values below 1 second might show better correlation with crashes than that observed for 1 second, particularly given that a threshold of 0 represents a crash, the data requirements with respect to the observation period increase substantially as the threshold falls below 1 second, as the event becomes increasingly rare. Svensson (1998) however observed that even though a large number of serious conflicts indicate unsafe conditions, frequent non-serious conflicts may also be an indication of safety. It alludes to the possibility of combining both tails of the PET distribution for better evaluation of overall safety. In such a case, the threshold and data requirement considerations might be quite different to that when using only one tail of the PET distribution. Therefore future research can work on this direction of exploring both tails of the PET distribution.

- The current research has only focused on PET and limited to the interactions between left-turn vehicles and opposing through vehicles. Only signalized intersections having a permitted left-turn phase were considered in this research. This research can be further extended to other types of conflicts, other types of locations or facilities, and other surrogates.

- There were only 25 respondents for the survey based on which the observations on expert evaluations were made in this thesis. A higher the number of respondents would increase the confidence in the observations and conclusions reached. Future research efforts could focus on increasing the number of responses to verify the observations made in this research. More importantly, future research should focus on the potential of the qualitative measures (based on human expert responses/opinions) to act as supplemental information to quantitative measures (surrogate measures, traffic characteristics etc.) in building better crash prediction models or in developing expert systems to identify hazardous locations.

APPENDIX A

CONDITIONAL PROBABILITIES OF PET AT DIFFERENT THRESHOLDS

Table A.1: Conditional Probabilities of PET at GA 138 and Sigman Rd (GA 20)

PET (x)	# of PETs <= x or 4-year crash count	Observed Cond. Prob. (Threshold)					
		4	3.5	3	2.5	2	1.5
4	68	1.0000					
3.5	51	0.7500	1.0000				
3	40	0.5882	0.7843	1.0000			
2.5	24	0.3529	0.4706	0.5853	1.0000		
2	13	0.1912	0.2549	0.3170	0.5416	1.0000	
1.5	9	0.1324	0.1765	0.2250	0.3750	0.6923	1.0000
1	8	0.1176	0.1569	0.1951	0.3330	0.6154	0.8889
0	90	0.0002	0.0002	0.0003	0.0004	0.0008	0.0012

Table A.2: Conditional Probabilities of PET at GA 20 and Willow Ln

PET (x)	# of PETs <= x or 4-year crash count	Observed Cond. Prob. (Threshold)					
		4	3.5	3	2.5	2	1.5
4	188	1.0000					
3.5	170	0.9043	1.0000				
3	121	0.6436	0.7118	1.0000			
2.5	98	0.5213	0.5765	0.8099	1.0000		
2	67	0.3564	0.3941	0.5537	0.6837	1.0000	
1.5	39	0.2074	0.2294	0.3223	0.3980	0.5821	1.0000
1	10	0.0532	0.0588	0.0826	0.1020	0.1493	0.2564
0	64	0.0001	0.0001	0.0001	0.0001	0.0001	0.0003

Table A.3: Conditional Probabilities of PET at N. Druid Hills Rd. and Lavista Rd.

PET (x)	# of PETs <= x or 4-year crash count	Observed Cond. Prob. (Threshold)					
		4	3.5	3	2.5	2	1.5
4	75	1.0000					
3.5	68	0.9067	1.0000				
3	61	0.8133	0.8971	1.0000			
2.5	45	0.6000	0.6618	0.7377	1.0000		
2	28	0.3733	0.4118	0.4590	0.6220	1.0000	
1.5	16	0.2133	0.2353	0.2623	0.3550	0.5714	1.0000
1	9	0.1200	0.1324	0.1475	0.2000	0.3214	0.5625
0	118	0.0002	0.0002	0.0003	0.0003	0.0005	0.0009

Table A.4: Conditional Probabilities of PET at Roswell Rd and Wieuca Rd

PET (x)	# of PETs <= x or 4-year crash count	Observed Cond. Prob. (Threshold)					
		4	**3.5**	3	**2.5**	**2**	**1.5**
4	121	1.0000					
3.5	97	0.8017	1.0000				
3	72	0.5950	0.7423	1.0000			
2.5	~~55~~	0.4545	0.5670	0.7639	1.0000		
2	29	0.2397	0.2990	0.4028	0.5273	1.0000	
1.5	14	0.1157	0.1443	0.1944	0.2545	0.4828	1.0000
1	8	0.0661	0.0825	0.1111	0.1455	0.2759	0.5714
0	78	0.0001	0.0002	0.0002	0.0002	0.0004	0.0009

Table A.5: Conditional Probabilities of PET at Grayson Hwy and Scenic Hwy

PET (x)	# of PETs <= x or 4-year crash count	Observed Cond. Prob. (Threshold)					
		4	**3.5**	3	**2.5**	**2**	**1.5**
4	123	1.0000					
3.5	105	0.8537	1.0000				
3	83	0.6748	0.7905	1.0000			
2.5	64	0.5203	0.6095	0.7711	1.0000		
2	38	0.3089	0.3619	0.4578	0.5938	1.0000	
1.5	23	0.1870	0.2190	0.2771	0.3594	0.6053	1.0000
1	6	0.0488	0.0571	0.0723	0.0938	0.1579	0.2609
0	53	0.0001	0.0001	0.0001	0.0001	0.0002	0.0004

Table A.6: Conditional Probabilities of PET at Lawrenceville Hwy and Lawrenceville Suwanee Rd

PET (x)	# of PETs <= x or 4-year crash count	Observed Cond. Prob. (Threshold)					
		4	**3.5**	3	**2.5**	**2**	**1.5**
4	83	0.99999					
3.5	64	0.771084	0.99999				
3	49	0.590361	0.765625	0.99999			
2.5	39	0.46988	0.609375	0.795918	0.99999		
2	19	0.228916	0.296875	0.387755	0.487179	0.99999	
1.5	9	0.108434	0.140625	0.183673	0.230769	0.473684	0.99999
1	2	0.024096	0.03125	0.040816	0.051282	0.105263	0.222222
0	73	0.000143	0.000186	0.000242	0.000304	0.000625	0.001319

Table A.7: Conditional Probabilities of PET at N. Druid Hills and Lawrenceville Hwy

PET (x)	# of PETs <= x or 4-year crash count	Observed Cond. Prob. (Threshold)					
		4	3.5	3	2.5	2	1.5
4	112	0.99999					
3.5	91	0.8125	0.99999				
3	71	0.633929	0.78022	0.99999			
2.5	44	0.392857	0.483516	0.619718	0.99999		
2	22	0.196429	0.241758	0.309859	0.5	0.99999	
1.5	12	0.107143	0.131868	0.169014	0.272727	0.545455	0.99999
1	1	0.008929	0.010989	0.014085	0.022727	0.045455	0.083333
0	48	6.73E-05	8.28E-05	0.000106	0.000171	0.000342	0.000628

Table A.8: Conditional Probabilities of PET at GA 10 and Grayson Pkwy

PET (x)	# of PETs <= x or 4-year crash count	Observed Cond. Prob. (Threshold)					
		4	3.5	3	2.5	2	1.5
4	112	0.99999					
3.5	91	0.839779	0.99999				
3	71	0.640884	0.763158	0.99999			
2.5	44	0.469613	0.559211	0.732759	0.99999		
2	22	0.243094	0.289474	0.37931	0.517647	0.99999	
1.5	12	0.082873	0.098684	0.12931	0.176471	0.340909	0.99999
1	1	0.01105	0.013158	0.017241	0.023529	0.045455	0.133333
0	48	3.53E-05	4.21E-05	5.51E-05	7.52E-05	0.000145	0.000426

Table A.9: Conditional Probabilities of PET at GA 10 and Oak Rd

PET (x)	# of PETs <= x or 4-year crash count	Observed Cond. Prob. (Threshold)					
		4	3.5	3	2.5	2	1.5
4	136	0.99999					
3.5	115	0.845588	0.99999				
3	96	0.705882	0.834783	0.99999			
2.5	58	0.426471	0.504348	0.604167	0.99999		
2	39	0.286765	0.33913	0.40625	0.672414	0.99999	
1.5	13	0.095588	0.113043	0.135417	0.224138	0.333333	0.99999
1	3	0.022059	0.026087	0.03125	0.051724	0.076923	0.230769
0	9	1.51E-05	1.79E-05	2.14E-05	3.54E-05	5.27E-05	0.000158

Table A.10: Conditional Probabilities of PET at Ponce De Leon Ave and Moreland Ave

PET (x)	# of PETs <= x or 4-year crash count	Observed Cond. Prob. (Threshold)					
		4	3.5	3	2.5	2	1.5
4	151	0.99999					
3.5	119	0.788079	0.99999				
3	94	0.622517	0.789916	0.99999			
2.5	73	0.483444	0.613445	0.776596	0.99999		
2	40	0.264901	0.336134	0.425532	0.547945	0.99999	
1.5	10	0.066225	0.084034	0.106383	0.136986	0.25	0.99999
1	3	0.019868	0.02521	0.031915	0.041096	0.075	0.3
0	27	2.72E-05	3.45E-05	4.37E-05	5.63E-05	0.000103	0.000411

Table A.11: Conditional Probabilities of PET at Memorial Dr and Columbia Dr

PET (x)	# of PETs <= x or 4-year crash count	Observed Cond. Prob. (Threshold)					
		4	3.5	3	2.5	2	1.5
4	93	0.99999					
3.5	71	0.763441	0.99999				
3	48	0.516129	0.676056	0.99999			
2.5	26	0.27957	0.366197	0.541667	0.99999		
2	11	0.11828	0.15493	0.229167	0.423077	0.99999	
1.5	4	0.043011	0.056338	0.083333	0.153846	0.363636	0.99999
1							
0	15	2.95E-05	3.86E-05	5.71E-05	0.000105	0.000249	0.000411

Table A.12: Conditional Probabilities of PET at Scott Blvd and Clairemont Ave

PET (x)	# of PETs <= x or 4-year crash count	Observed Cond. Prob. (Threshold)					
		4	3.5	3	2.5	2	1.5
4	107	0.99999					
3.5	76	0.71028	0.99999				
3	66	0.616822	0.868421	0.99999			
2.5	53	0.495327	0.697368	0.80303	0.99999		
2	32	0.299065	0.421053	0.484848	0.603774	0.99999	
1.5	11	0.102804	0.144737	0.166667	0.207547	0.34375	0.99999
1	5	0.046729	0.065789	0.075758	0.09434	0.15625	0.454545
0	23	3.2E-05	4.51E-05	5.19E-05	6.46E-05	0.000107	0.000311

Table A.13: Conditional Probabilities of PET at Glenwood Rd and Columbia Dr

PET (x)	# of PETs <= x or 4-year crash count	Observed Cond. Prob. (Threshold)					
		4	**3.5**	3	**2.5**	**2**	**1.5**
4	106	0.99999					
3.5	83	0.783019	0.99999				
3	59	0.556604	0.710843	0.99999			
2.5	32	0.301887	0.385542	0.542373	0.99999		
2	14	0.132075	0.168675	0.237288	0.4375	0.99999	
1.5	5	0.04717	0.060241	0.084746	0.15625	0.357143	0.99999
1	1	0.009434	0.012048	0.016949	0.03125	0.071429	0.2
0	15	2.15E-06	2.75E-06	3.87E-06	7.13E-06	1.63E-05	4.57E-05

Table A.14: Conditional Probabilities of PET at Buford Hwy and Sugarloaf Pkwy

PET (x)	# of PETs <= x or 4-year crash count	Observed Cond. Prob. (Threshold)					
		4	**3.5**	3	**2.5**	**2**	**1.5**
4	51	0.99999					
3.5	36	0.705882	0.99999				
3	22	0.431373	0.611111	0.99999			
2.5	17	0.333333	0.472222	0.772727	0.99999		
2	9	0.176471	0.25	0.409091	0.529412	0.99999	
1.5	4	0.078431	0.111111	0.181818	0.235294	0.444444	0.99999
1							
0	6	8.95E-06	1.27E-05	2.08E-05	2.69E-05	5.07E-05	0.000114

Table A.15: Conditional Probabilities of PET at MLK Jr Blvd and Brownlee Rd

PET (x)	# of PETs <= x or 4-year crash count	Observed Cond. Prob. (Threshold)					
		4	**3.5**	3	**2.5**	**2**	**1.5**
4	51	40	0.99999				
3.5	36	34	0.85	0.99999			
3	22	27	0.675	0.794118	0.99999		
2.5	17	17	0.425	0.5	0.62963	0.99999	
2	9	8	0.2	0.235294	0.296296	0.470588	0.99999
1.5	4	2	0.05	0.058824	0.074074	0.117647	0.25
1							
0	2	1.14E-05	1.34E-05	1.69E-05	2.69E-05	5.71E-05	0.000114

Table A.16: Conditional Probabilities of PET at Whitlock Ave and Lindley Ave

PET (x)	# of PETs <= x or 4-year crash count	Observed Cond. Prob. (Threshold)					
		4	3.5	3	2.5	2	1.5
4	40	0.99999					
3.5	32	0.8	0.99999				
3	26	0.65	0.8125	0.99999			
2.5	20	0.5	0.625	0.769231	0.99999		
2	12	0.3	0.375	0.461538	0.6	0.99999	
1.5	1	0.025	0.03125	0.038462	0.05	0.083333	
1							
0	2	5.71E-06	7.13E-06	8.78E-06	1.14E-05	1.9E-05	

Table A.17: Conditional Probabilities of PET at North Ave and Techwood Dr

PET (x)	# of PETs <= x or 4-year crash count	Observed Cond. Prob. (Threshold)					
		4	3.5	3	2.5	2	1.5
4	37	0.99999					
3.5	29	0.783784	0.99999				
3	21	0.567568	0.724138	0.99999			
2.5	15	0.405405	0.517241	0.714286	0.99999		
2	8	0.216216	0.275862	0.380952	0.533333	0.99999	
1.5	3	0.081081	0.103448	0.142857	0.2	0.375	
1							
0	2	6.17E-06	7.87E-06	1.09E-05	1.52E-05	2.85E-05	

Table A.18: Conditional Probabilities of PET at Cobb Pkwy and Gresham Rd

PET (x)	# of PETs <= x or 4-year crash count	Observed Cond. Prob. (Threshold)					
		4	3.5	3	2.5	2	1.5
4	116	0.99999					
3.5	94	0.810345	0.99999				
3	60	0.517241	0.638298	0.99999			
2.5	48	0.413793	0.510638	0.8	0.99999		
2	28	0.241379	0.297872	0.466667	0.583333	0.99999	
1.5	15	0.12931	0.159574	0.25	0.3125	0.535714	0.99999
1	2	0.017241	0.021277	0.033333	0.041667	0.071429	0.133333
0	5	3.94E-06	4.86E-06	7.61E-06	9.51E-06	1.63E-05	3.04E-05

APPENDIX B

OBSERVED AND GEV FITTED CDF VALUES FOR PET DATA AT STUDY INTERSECTIONS

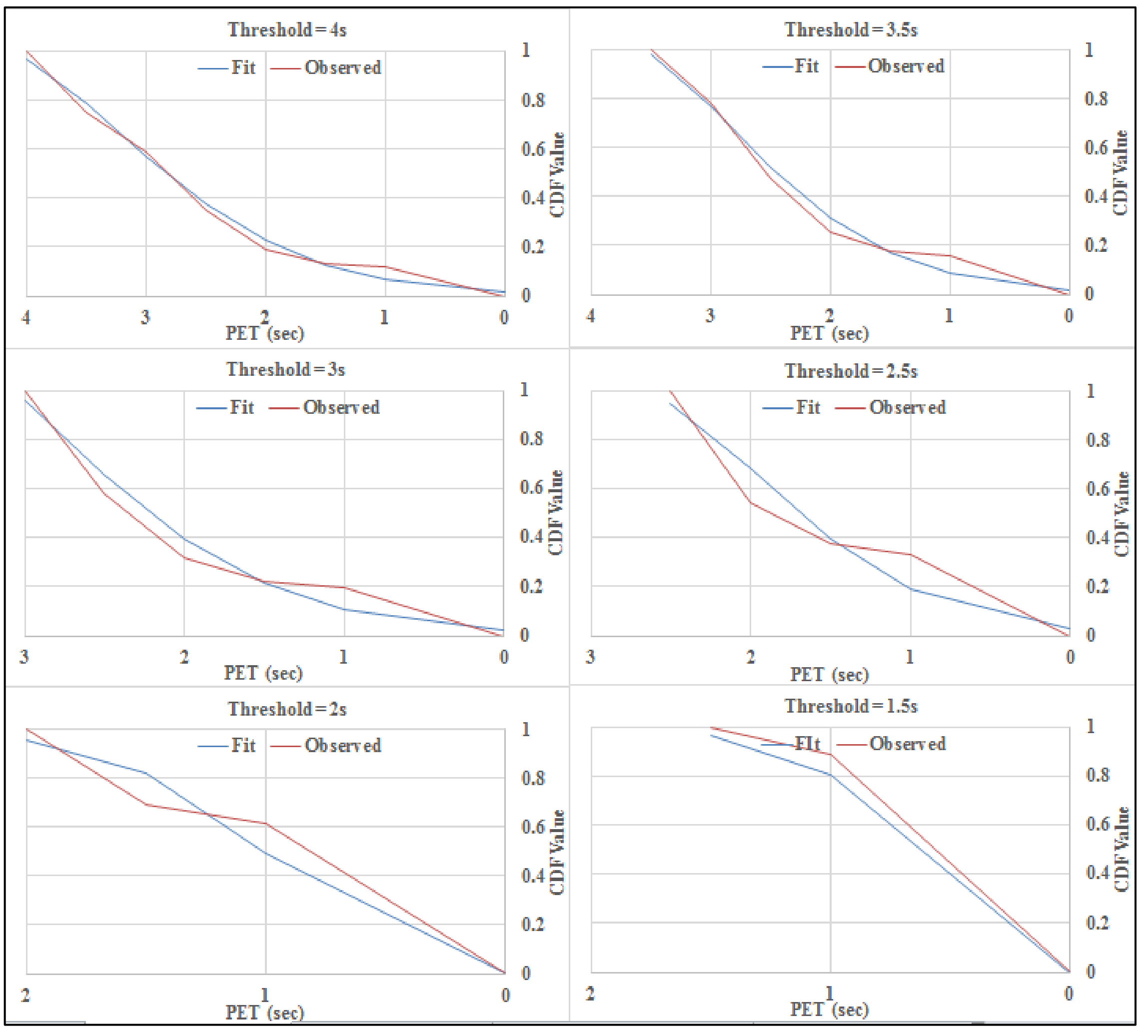


Figure B.1: Observed and GEV fitted CDF values for PET data at GA 138 and Sigman Rd (GA 20)

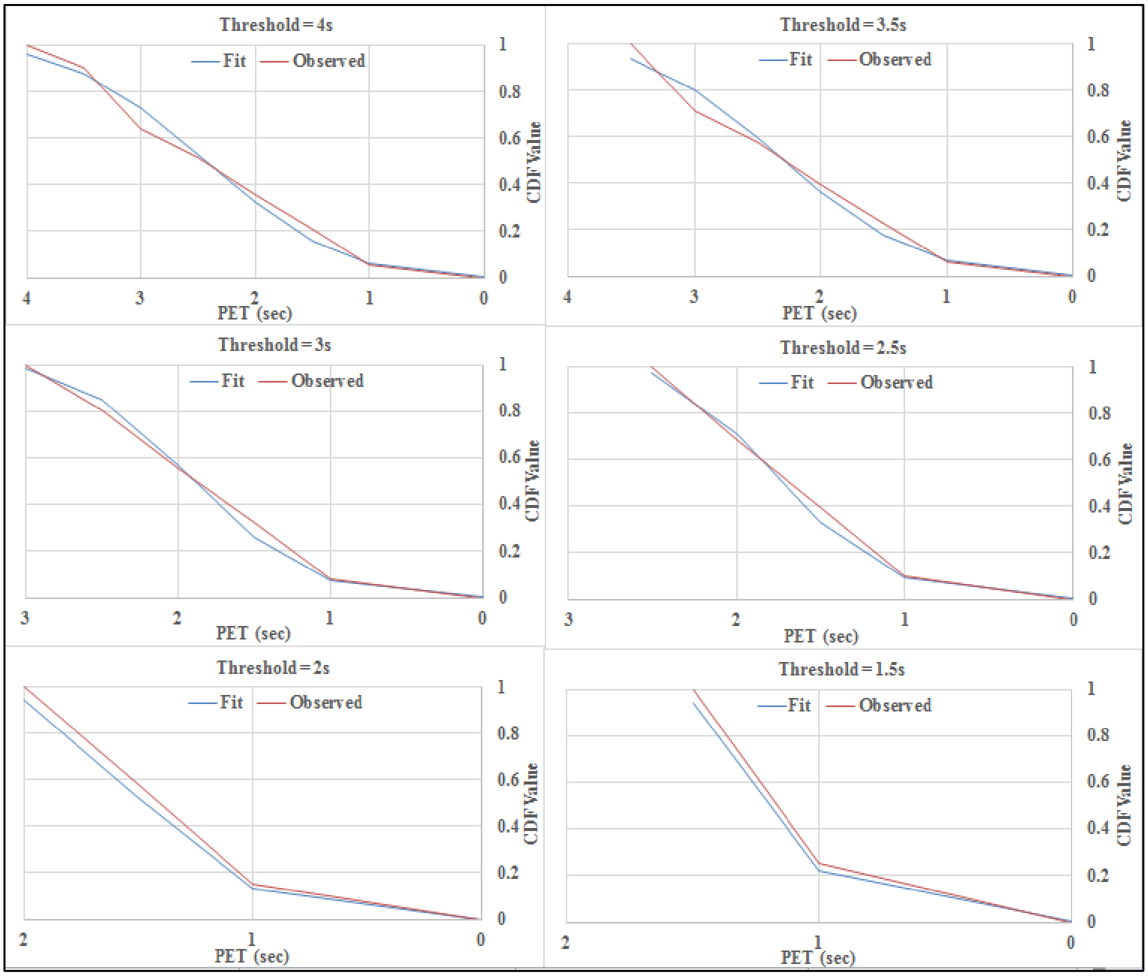


Figure B.2: Observed and GEV fitted CDF values for PET data at GA 20 and Willow Ln

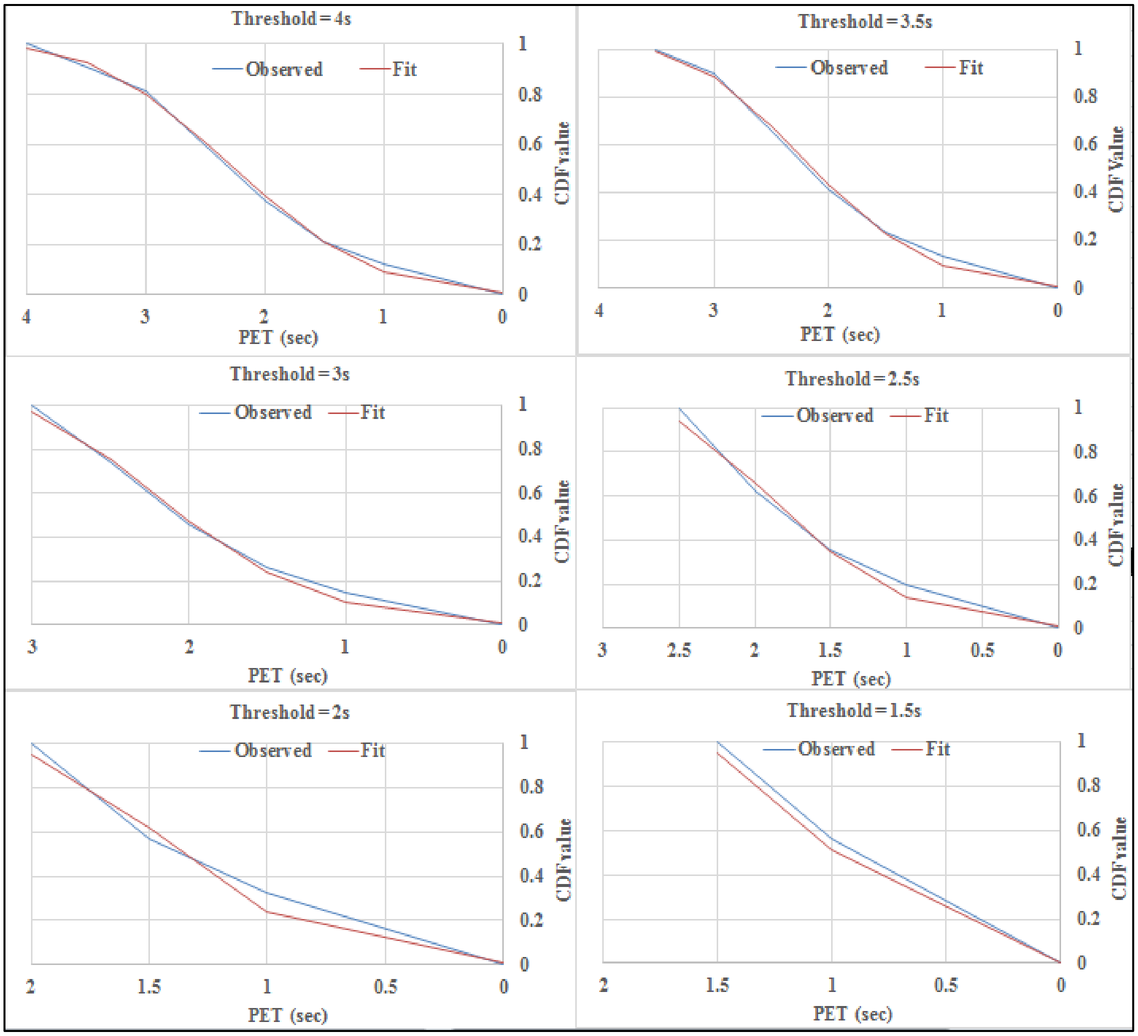


Figure B.3: Observed and GEV fitted CDF values for PET data at N. Druid Hills Rd and Lavista Rd.

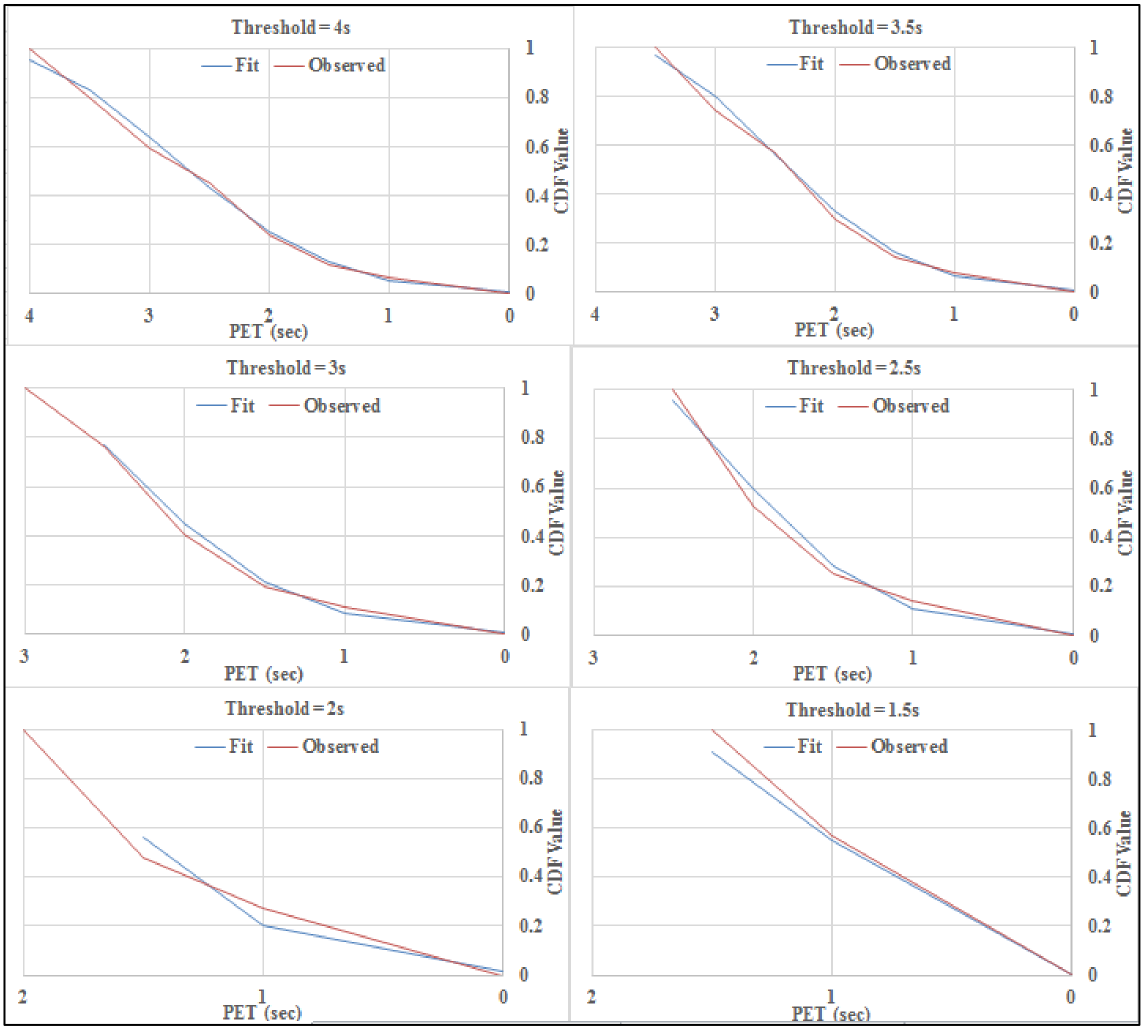


Figure B.4: Observed and GEV fitted CDF values for PET data at Roswell Rd and Wieuca Rd

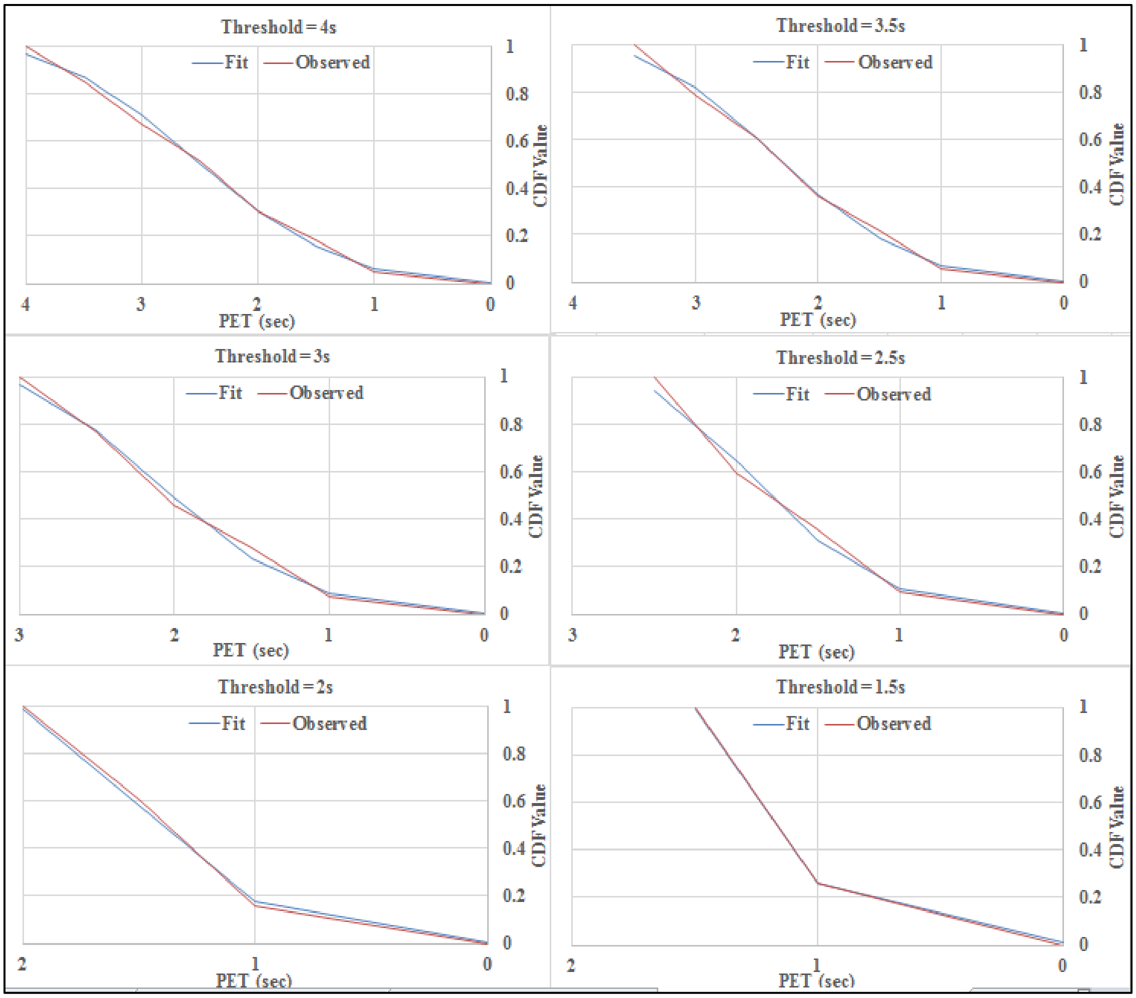


Figure B.5: Observed and GEV fitted CDF values for PET data at Grayson Hwy and Scenic Hwy

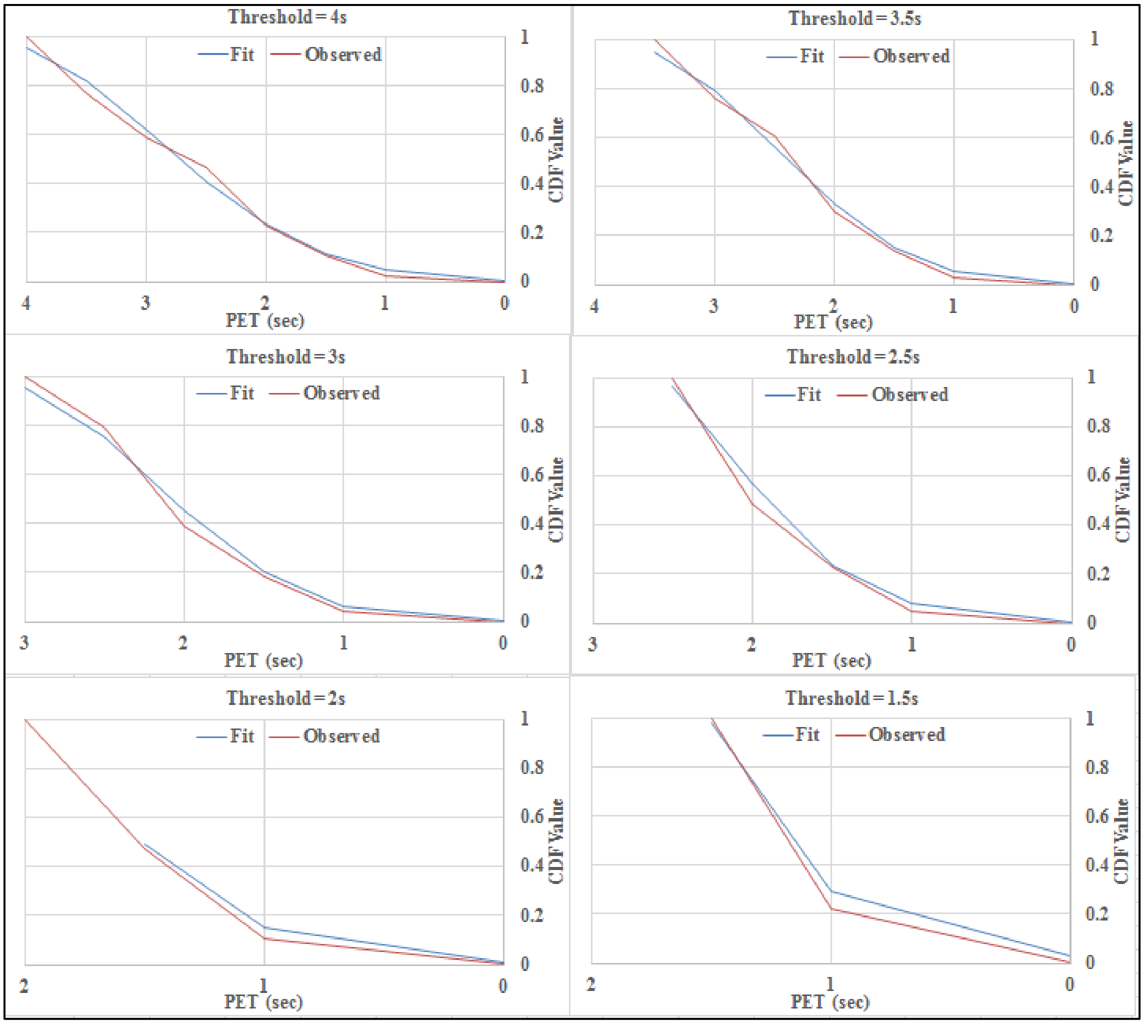


Figure B.6: Observed and GEV fitted CDF values for PET data at Lawrenceville Hwy and Lawrenceville Suwanee Rd.

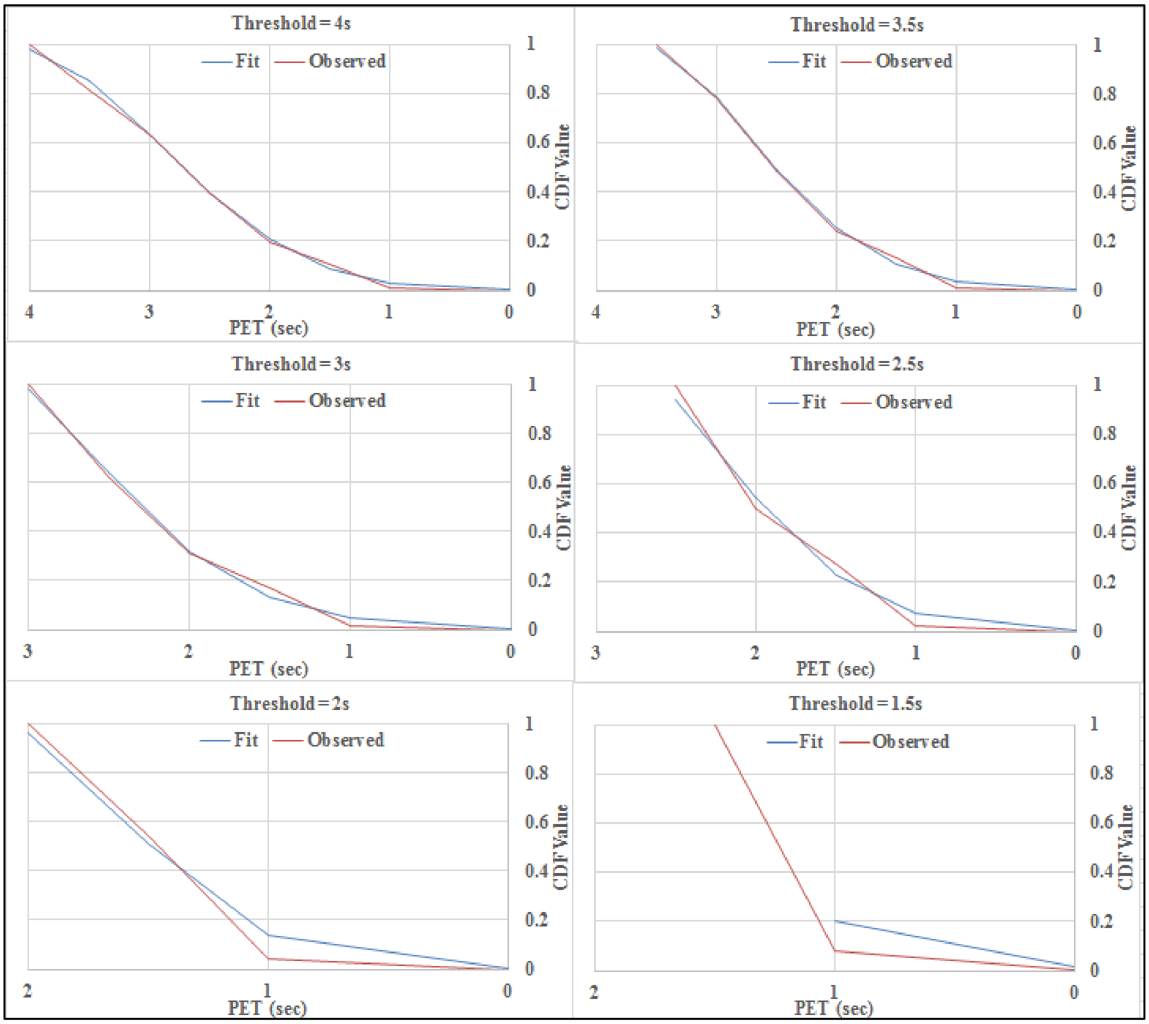


Figure B.7: Observed and GEV fitted CDF values for PET data at N. Druid Hills Rd. and Lawrenceville Hwy.

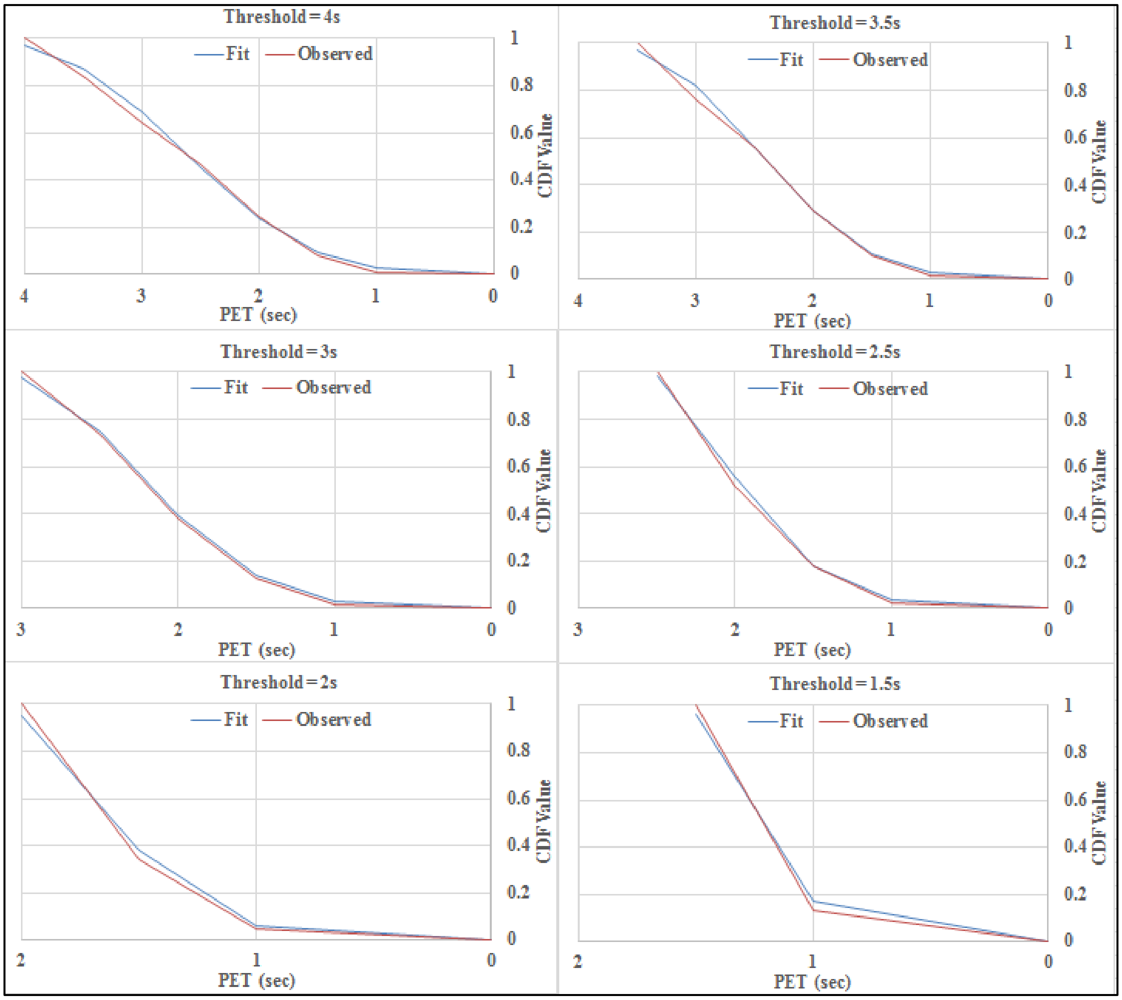


Figure B.8: Observed and GEV fitted CDF values for PET data at GA 10 and Grayson Pkwy

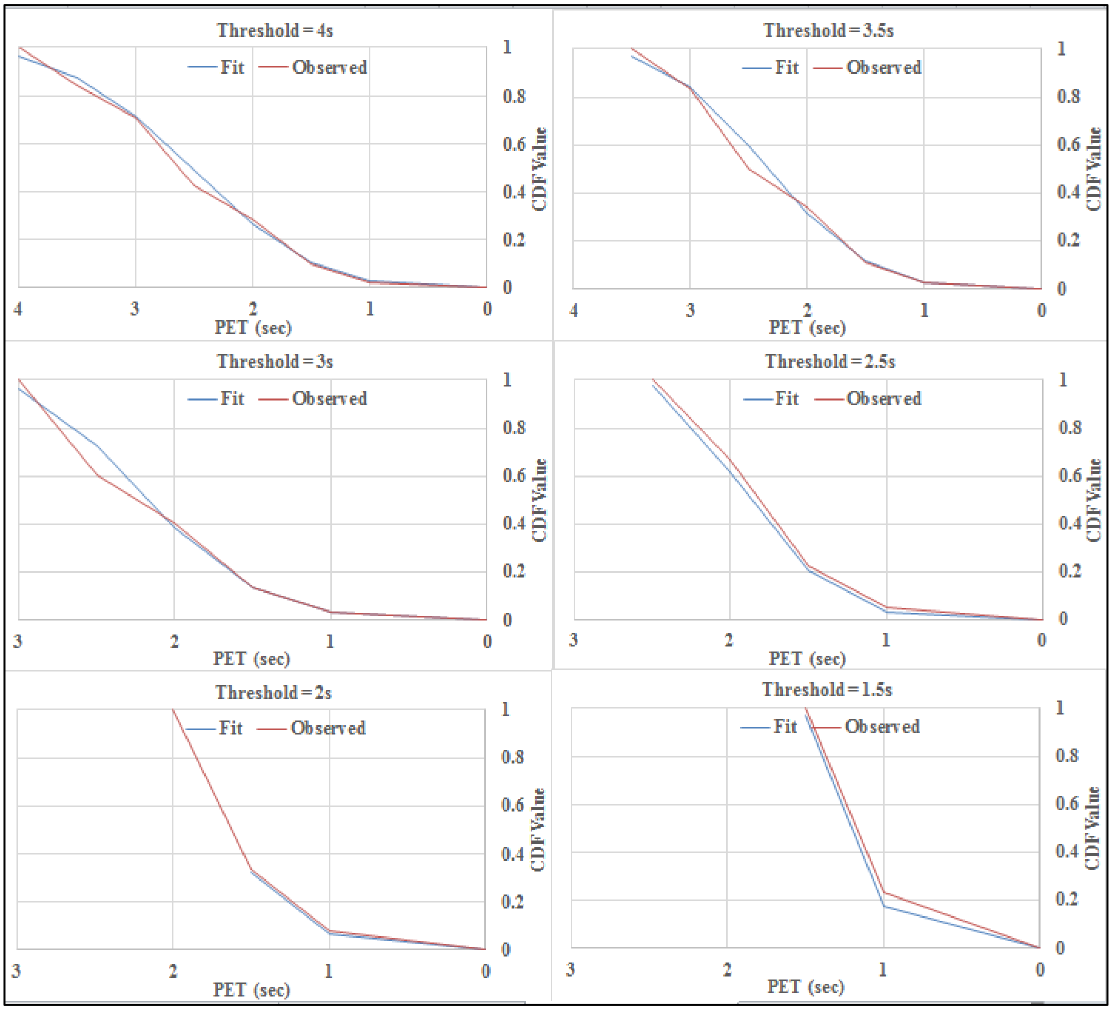


Figure B.9: Observed and GEV fitted CDF values for PET data at GA 10 and Oak Rd

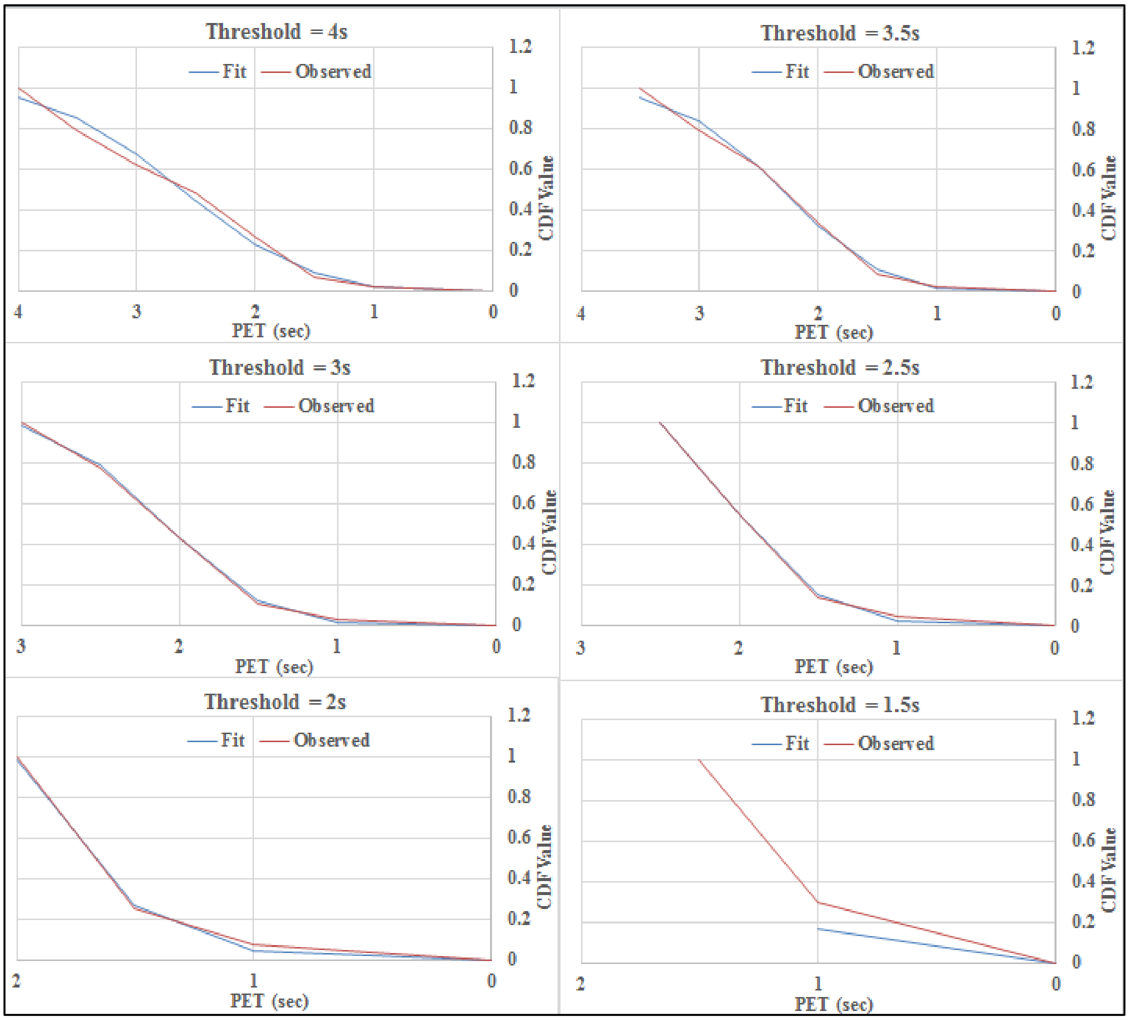


Figure B.10: Observed and GEV fitted CDF values for PET data at Ponce De Leon Ave and Moreland Ave

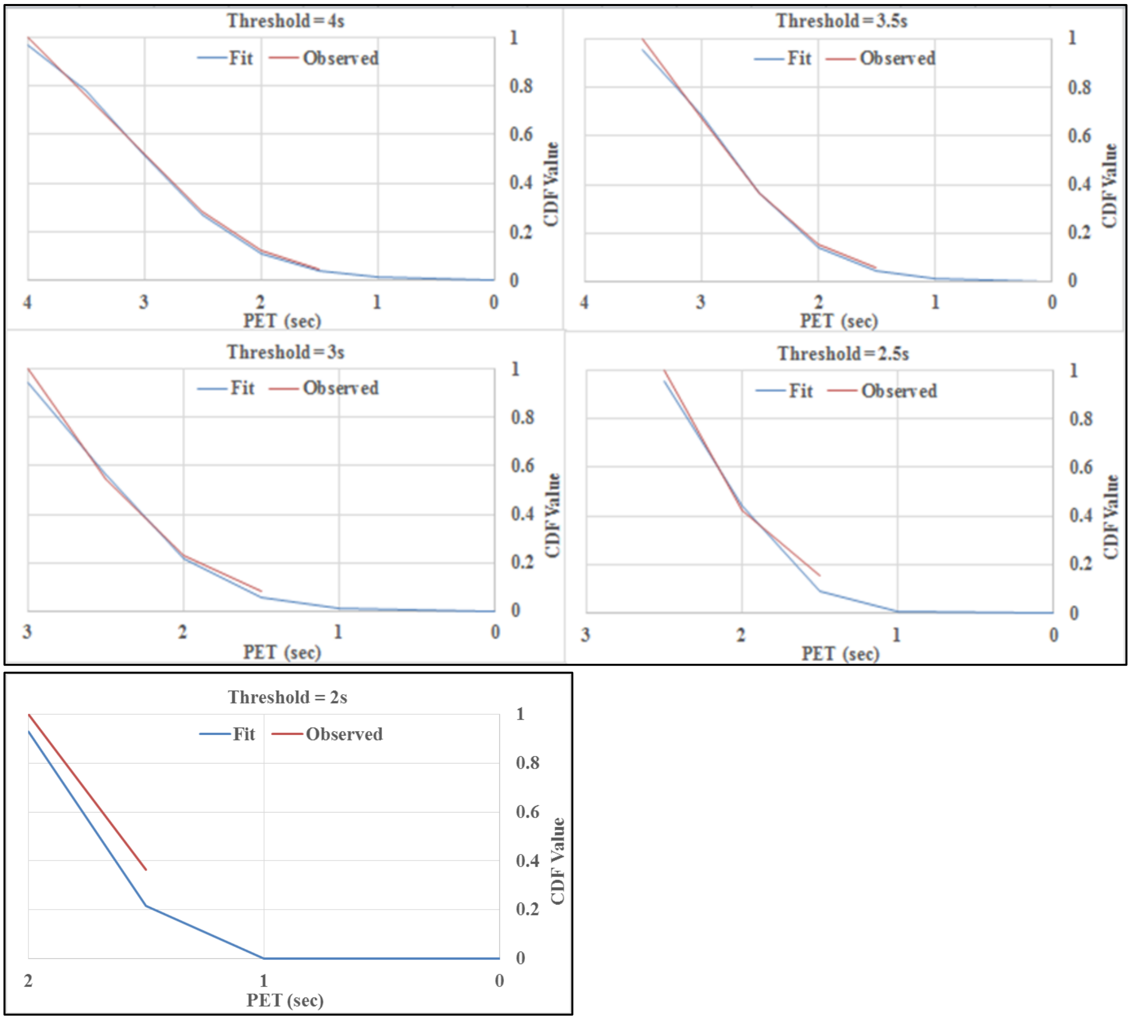


Figure B.11: Observed and GEV fitted CDF values for PET data at Memorial Dr and Covington Hwy

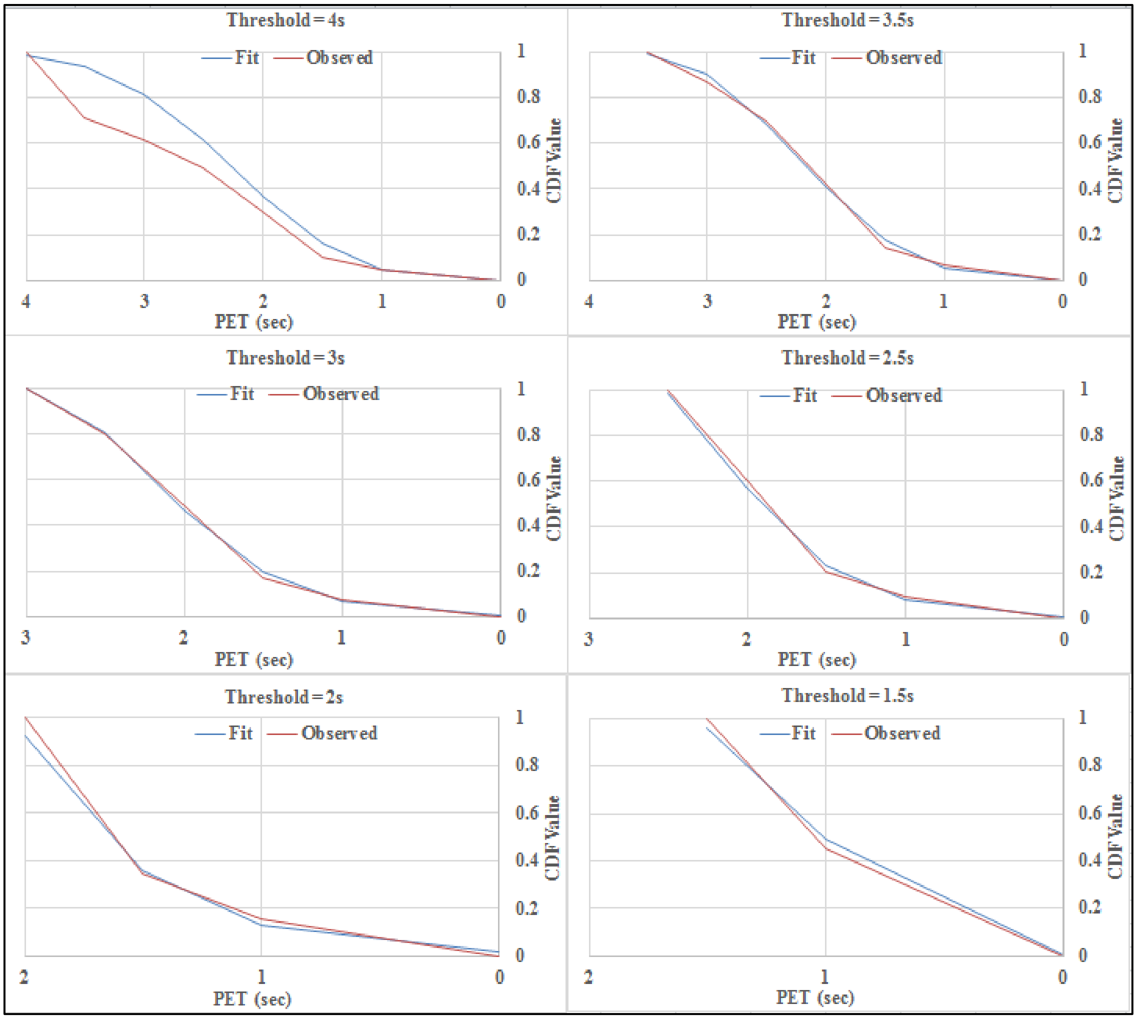


Figure B.12: Observed and GEV fitted CDF values for PET data at Scott Blvd and Clairemont Ave

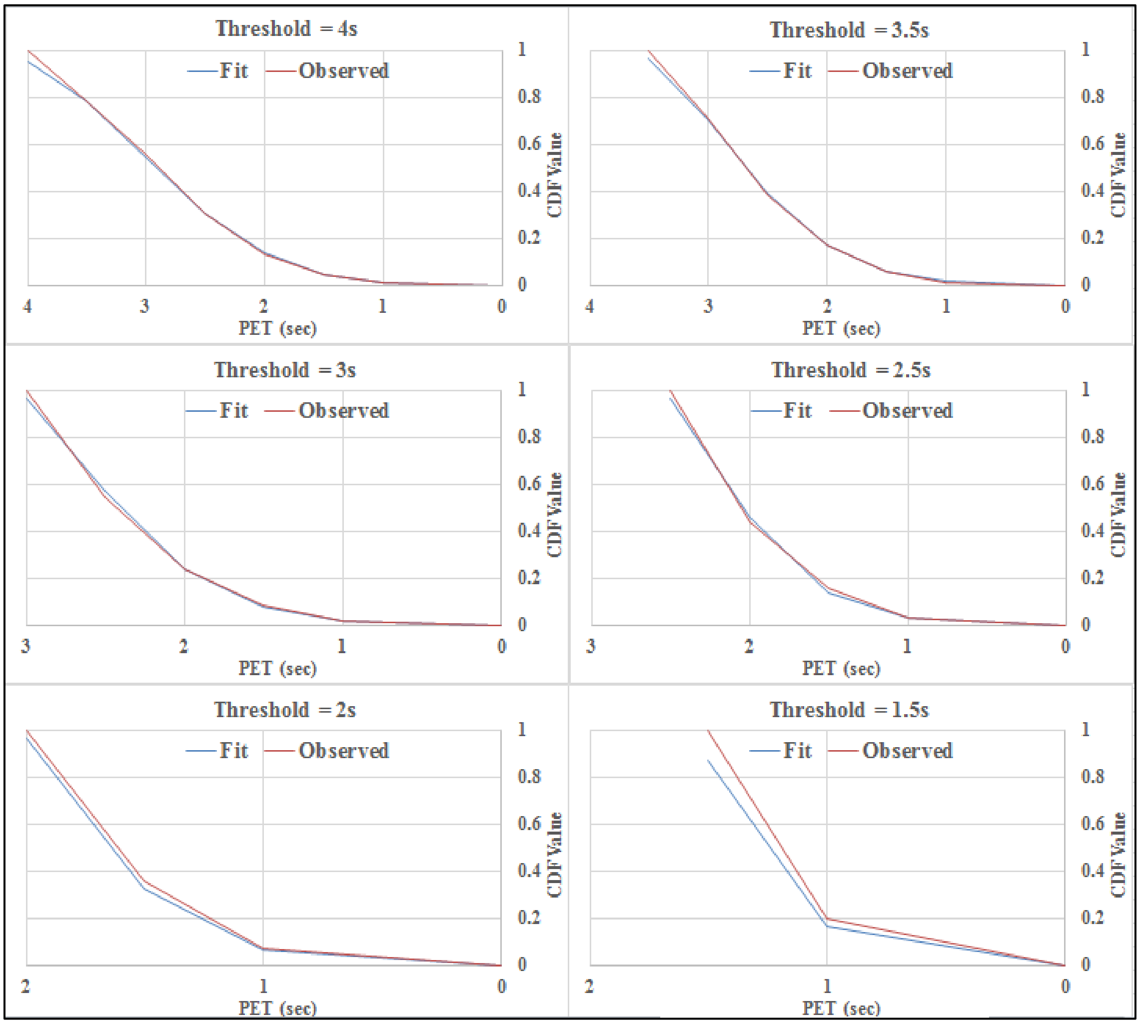


Figure B.13: Observed and GEV fitted CDF values for PET data at Glenwood Rd and Columbia Dr

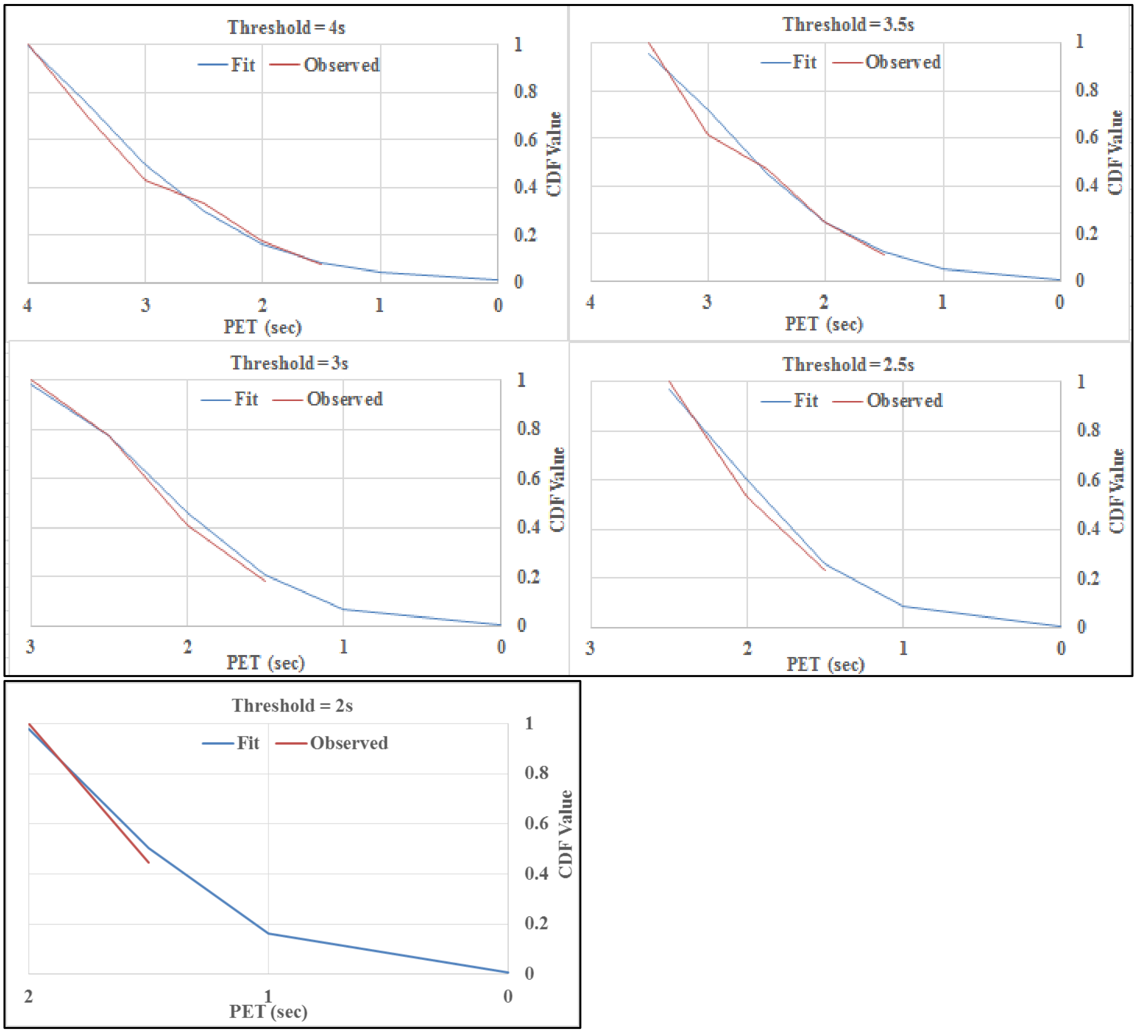


Figure B.14: Observed and GEV fitted CDF values for PET data at Buford Hwy and Sugarloaf Pkwy

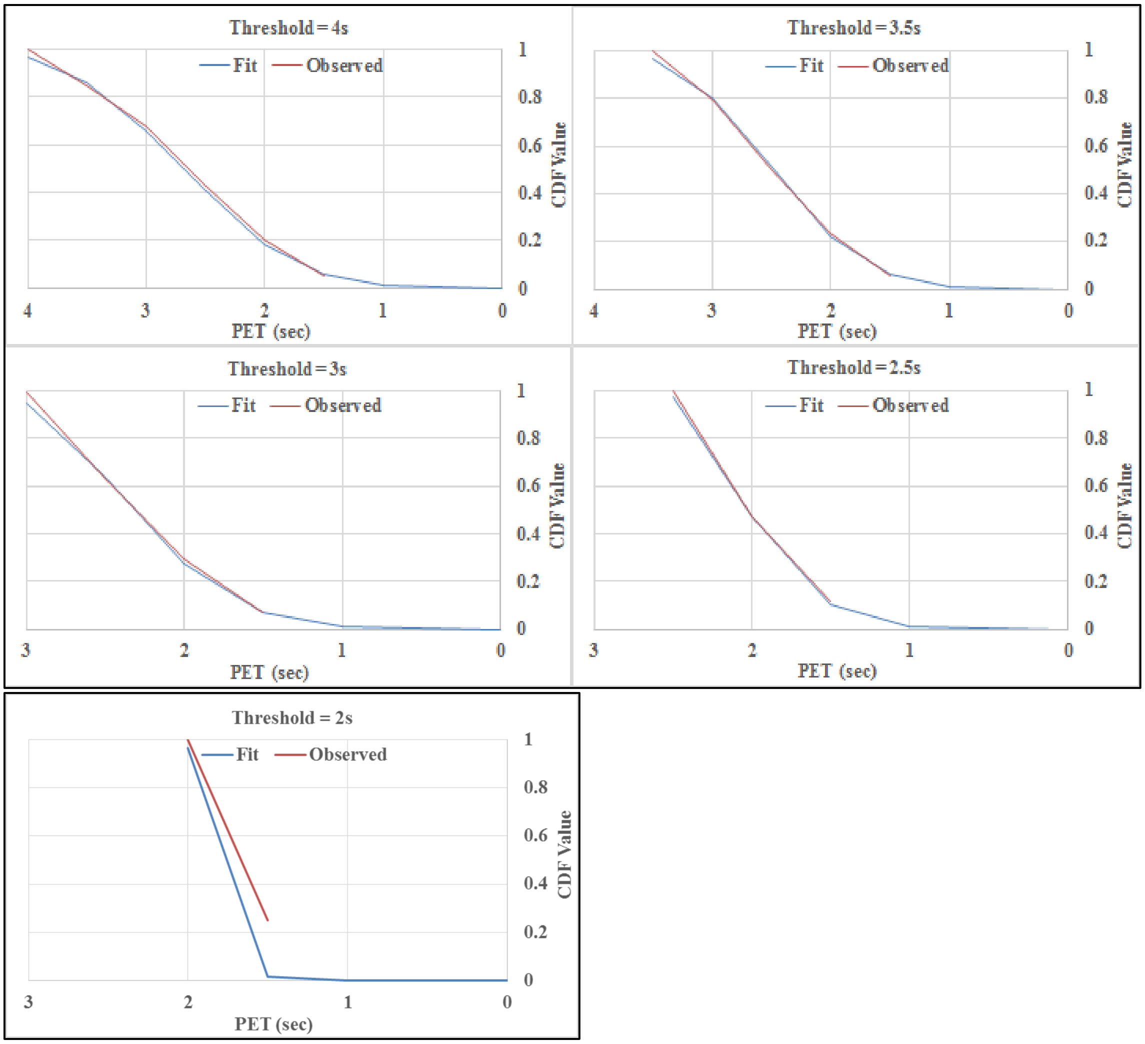


Figure B.15: Observed and GEV fitted CDF values for PET data at MLK Jr Blvd and Brownlee Rd

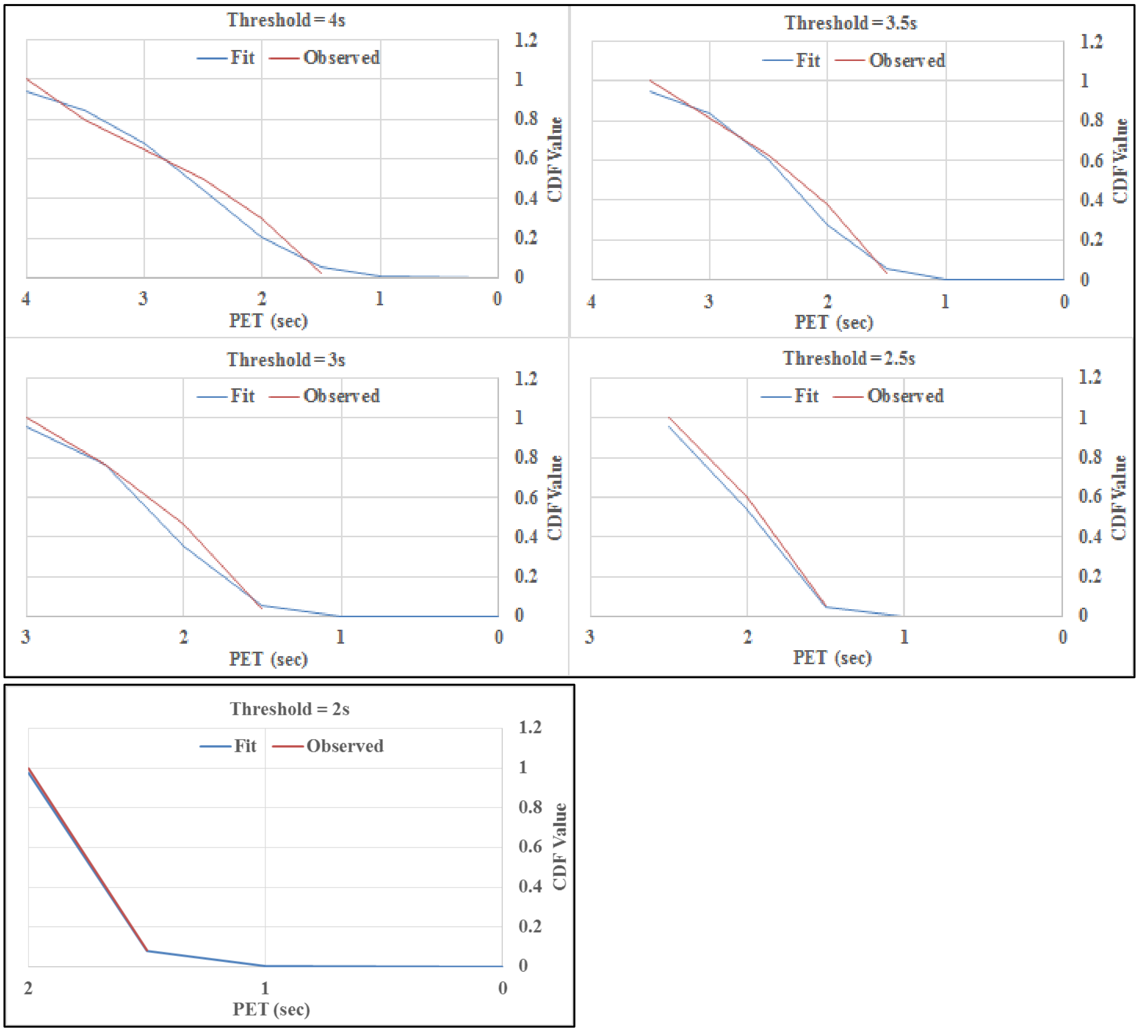


Figure B.16: Observed and GEV fitted CDF values for PET data at Whitlock Ave and Lindley Ave

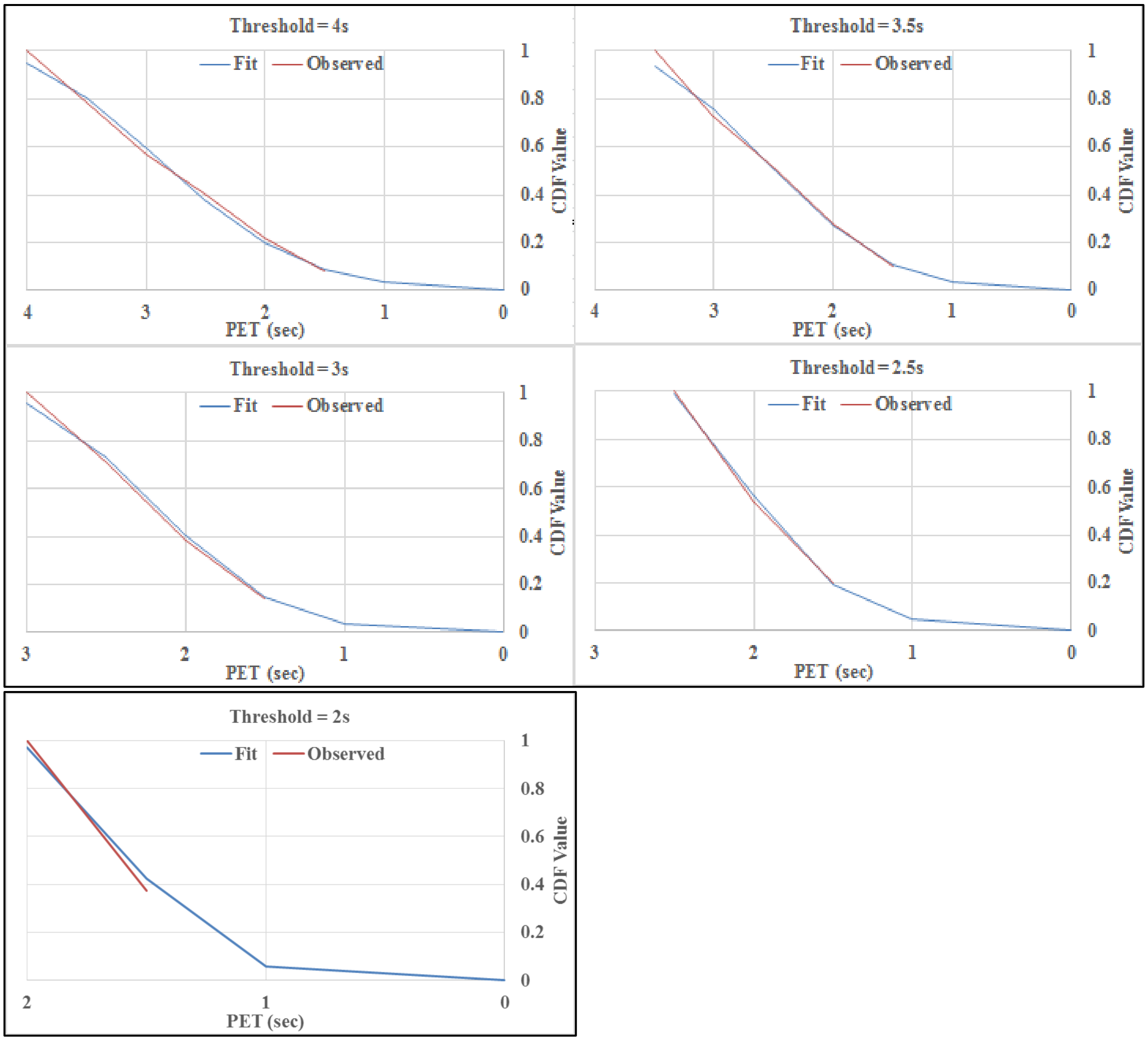


Figure B.17: Observed and GEV fitted CDF values for PET data at North Ave and Techwood Dr.

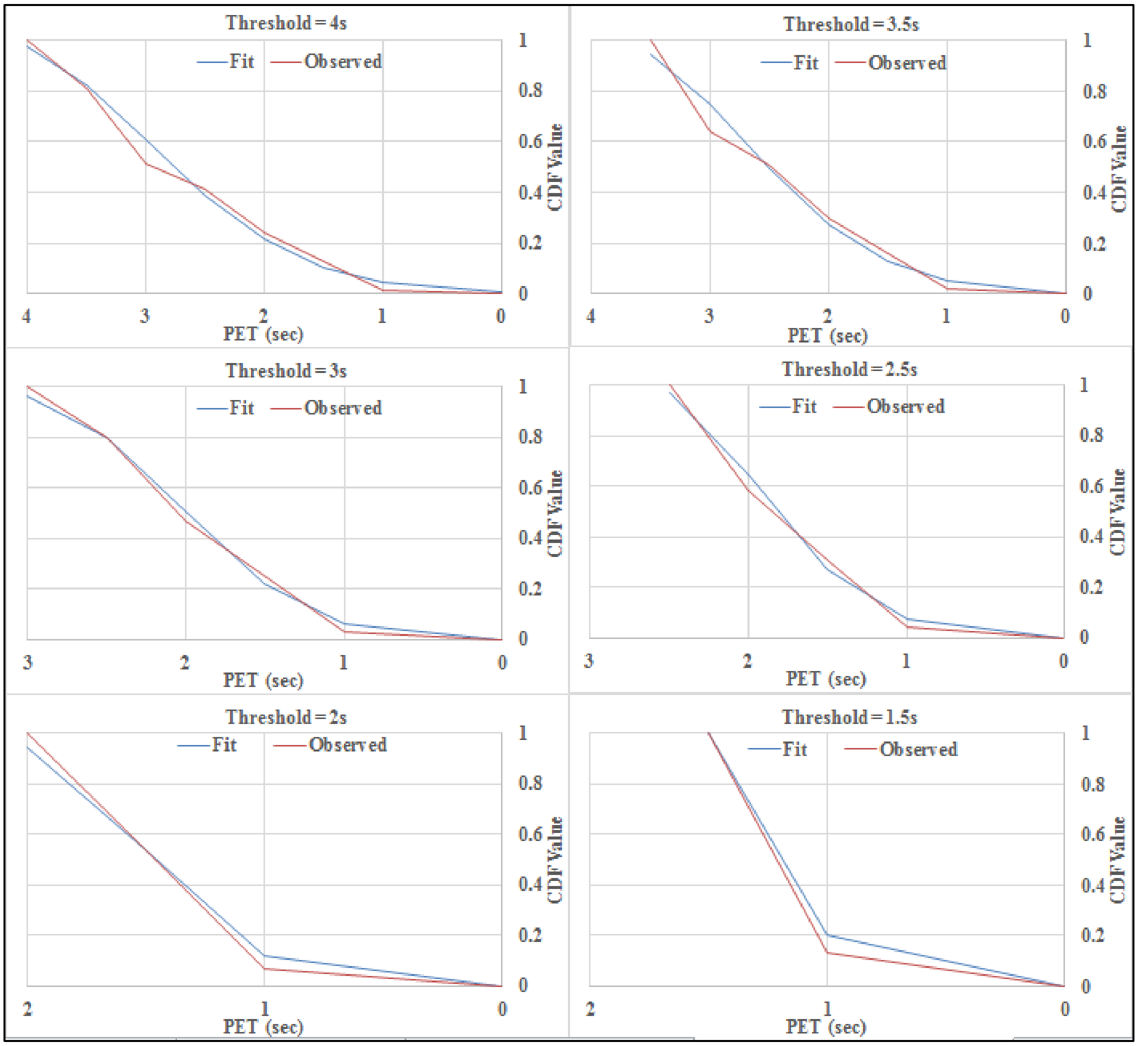


Figure B.18: Observed and GEV fitted CDF values for PET data at Cobb Pkwy and Gresham Rd

APPENDIX C

METADATA OF INTERSECTIONS USED IN THE SURVEY

Cobb Pkwy. and Gresham Rd.

Main Street AADT = 29820
Minor Street AADT = 6500
Avg. grade on major road = 1.8 degrees
Avg. grade on minor road = 0.5 degrees
Maximum avg. approach grade = 3.1 degrees

GA 10 and Grayson Pkwy.

Main Street AADT = 41400
Minor Street AADT = 10853
Avg. grade on major road = 2.3 degrees
Avg. grade on minor road = 0.7 degrees
Maximum avg. approach grade = 2.6 degrees

GA 10 and Henry Clower Blvd/Oak Rd.

Main Street AADT = 33630
Minor Street AADT = 2500
Avg. grade on major road = 0.7 degrees
Avg. grade on minor road = 0.6 degrees
Maximum avg. approach grade = 1.3 degrees

GA 20 and Willow Ln.

Main Street AADT = 23380
Minor Street AADT = 8000
Avg. grade on major road = 2.5 degrees
Avg. grade on minor road = 1.2 degrees
Maximum avg. approach grade = 2.6 degrees

GA 138 and Sigman Rd (GA 20)

Main Street AADT = 25655
Minor Street AADT = 19060
Avg. grade on major road = 1.4 degrees
Avg. grade on minor road = 0.8 degrees
Maximum avg. approach grade = 2.5 degrees

Glenwood Rd. - Columbia Dr.

Main Street AADT = 18360
Minor Street AADT = 17025
Avg. grade on major road = 1.0 degrees
Avg. grade on minor road = 2.1 degrees
Maximum avg. approach grade = 2.7 degrees

Grayson Hwy and Scenic Hwy

Main Street AADT = 30928
Minor Street AADT = 25918
Avg. grade on major road = 0.9 degrees
Avg. grade on minor road = 1.6 degrees
Maximum avg. approach grade = 2.1 degrees

Lawrenceville Hwy. and Lawrenceville Suwanee Rd.

Main Street AADT = 25594
Minor Street AADT = 17896
Avg. grade on major road = 1.6 degrees
Avg. grade on minor road = 0.6 degrees
Maximum avg. approach grade = 2.5 degrees

Memorial Dr. and Covington Hwy.

Main Street AADT = 25180
Minor Street AADT = 14365
Avg. grade on major road = 1.1 degrees
Avg. grade on minor road = 0.4 degrees
Maximum avg. approach grade = 1.4 degrees

MLK Jr Blvd. and Brownlee Rd.

Main Street AADT = 23360
Minor Street AADT = 1000
Avg. grade on major road = 1.8 degrees
Avg. grade on minor road = 2.3 degrees
Maximum avg. approach grade = 4.1 degrees

N Druid Hills Rd. and Lavista Rd.

Main Street AADT = 33600
Minor Street AADT = 15205
Avg. grade on major road = 1.2 degrees
Avg. grade on minor road = 2.7 degrees
Maximum avg. approach grade = 3.8 degrees

N Druid Hills Rd. and Lawrenceville Hwy.

Main Street AADT = 27415
Minor Street AADT = 22280
Avg. grade on major road = 1.2 degrees
Avg. grade on minor road = 1.9 degrees
Maximum avg. approach grade = 2.0 degrees

North Ave and Techwood Dr.

Main Street AADT = 20240
Minor Street AADT = 3470
Avg. grade on major road = 1.1 degrees
Avg. grade on minor road = 2.1 degrees
Maximum avg. approach grade = 2.4 degrees

Ponce De Leon Ave. and Moreland Ave.

Main Street AADT = 32320
Minor Street AADT = 22220
Avg. grade on major road = 1.3 degrees
Avg. grade on minor road = 1.7 degrees
Maximum avg. approach grade = 2 degrees

Roswell Rd. and W. Wieuca Rd.

Main Street AADT = 30180
Minor Street AADT = 8020
Avg. grade on major road = 0.9 degrees
Avg. grade on minor road = 1.2 degrees
Maximum avg. approach grade = 1.5 degrees

Scott Blvd. and Clairemont Ave.

Main Street AADT = 34925
Minor Street AADT = 23725
Avg. grade on major road = 2.8 degrees
Avg. grade on minor road = 0.9 degrees
Maximum avg. approach grade = 3.0 degrees

Sugarloaf Pkwy. and Buford Hwy.

Main Street AADT = 26210
Minor Street AADT = 10500
Avg. grade on major road = 1.0 degrees
Avg. grade on minor road = 0.3 degrees
Maximum avg. approach grade = 1.6 degrees

Whitlock Ave. and Lindley Ave.

Main Street AADT = 27230
Minor Street AADT = 6000
Avg. grade on major road = 0.9 degrees
Avg. grade on minor road = 0.6 degrees
Maximum avg. approach grade = 1.2 degrees

APPENDIX D

PICTURES OF STUDY INTERSECTIONS USED IN THE SURVEY

All Figures used in this appendix are courtesy (Peesapati, 2013)

Figure D.1: Lawrenceville Hwy. and Lawrenceville Suwanee Rd.

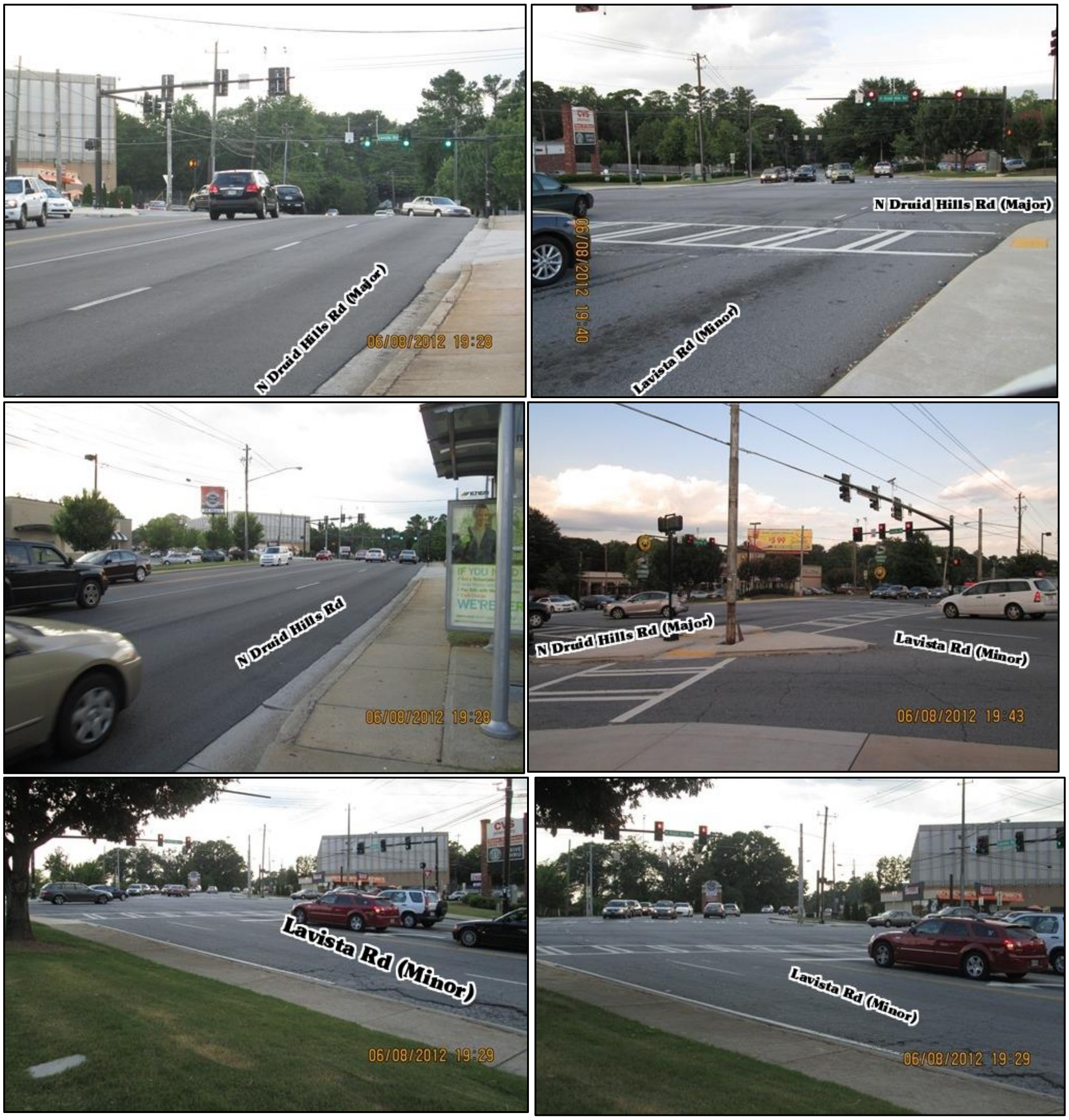


Figure D.2: N Druid Hills Rd. and Lavista Rd.

Figure D.3: Roswell Rd. and W Wieuca Rd.

Figure E.4: Grayson Hwy. and Scenic Hwy.

Figure D.5: Willow Ln. and SR 20

Figure D.6: Georgia SR 138 and Georgia SR 20

Figure D.7: Georgia SR 10 and Grayson Pkwy.

Figure D.8: Georgia SR 10 and Oak Rd/Henry Clower Blvd.

Figure D.9: Memorial Dr. and Covington Hwy.

Figure D.10: N. Druid Hills Rd. and Lawrenceville Hwy.

Figure D.11: Ponce De Leon Ave. and Moreland Ave.

Figure D.12: Scott Blvd. and Clairemont Ave.

Figure D.13: Buford Hwy. and Sugarloaf Pkwy

Figure D.14: Cobb Pkwy and Gresham Rd.

Figure D.15: Whitlock Ave. and Lindley Ave.

Figure D.16: North Ave. and Techwood Dr.

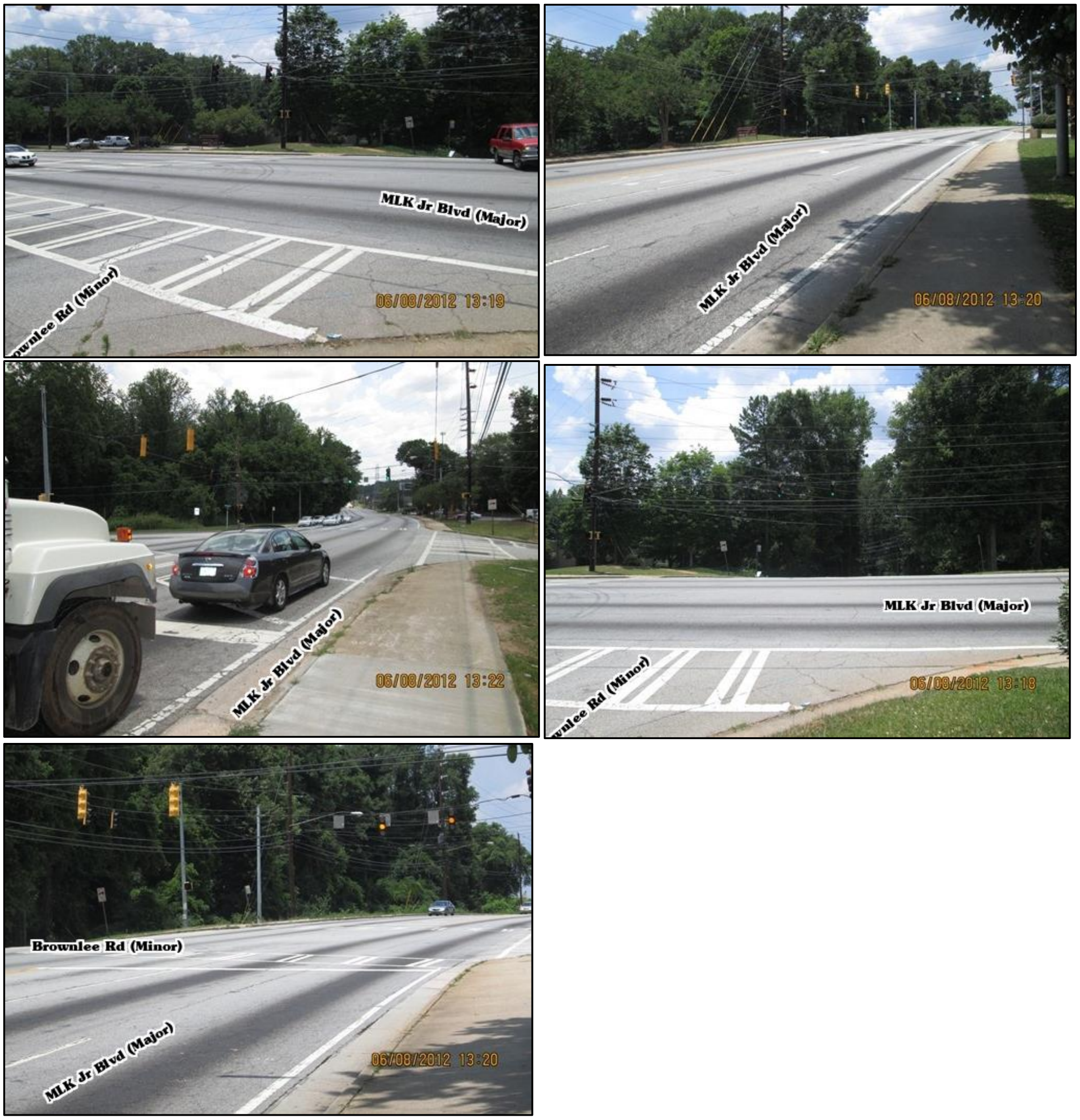


Figure D.17: MLK Jr. Blvd. and Brownlee Rd.

Figure D.18: Glenwood Rd. and Columbia Dr.

APPENDIX E

RELATIVE RANKINGS FOR INTERSECTIONS GIVEN BY EXPERTS

The Tables in this appendix show a group of four abbreviated intersections that the experts relatively ranked. The rankings based on crash numbers are also shown for comparison purposes.

The abbreviations of intersection names used in this appendix are as follows:

NDH_Lav – N. Druid Hills Rd. and Lavista Rd

GA138_Sig – GA 138 and Sigman Rd (GA 20)

Gra_Sce – Grayson Hwy and Scenic Hwy

GA20_Wil – GA 20 and Willow Ln

Law_Suw – Lawrenceville Hwy and Lawrenceville Suwanee Rd

Ros_Wie – Roswell Rd and Wieuca Rd

GA10_Gra – GA 10 and Grayson Pkwy

GA10_Oak – GA 10 and Oak Rd

Sco_Cla – Scott Blvd and Clairemont Ave

Pon_Mor – Ponce De Leon Ave and Moreland Ave

NDH_Law – N. Druid Hills Rd and Lawrenceville Hwy

Gle_Col – Glenwood Rd and Columbia Dr

Sug_Buf – Sugarloaf Pkwy and Buford Hwy

Mem_Cov – Memorial Dr and Covington Hwy

MLK_Bro – MLK Jr Blvd and Brownlee Rd

Whi_Lin – Whitlock Ave and Lindley Ave

Nor_Tec – North Ave and Techwood Dr

Cob_Gre – Cobb Pkwy and Gresham Rd

Table E.1: Relative ranking by experts vs. based on crash frequency

	Gra_Sce	Pon_Mor	Sco_Cla	Sug_Buf
Crash	4	1	2	3
Expert	1	2	3	4
	Gle_Col	**NDH_Law**	**Nor_Tec**	**Sug_Buf**
Crash	1	3	2	4
Expert	2	1	4	3
	GA10_Oak	**Law_Suw**	**NDH_Lav**	**NDH_Law**
Crash	3	1	2	4
Expert	4	2	1	3
	GA10_Gra	**GA138_Sig**	**Law_Suw**	**Ros_Wie**
Crash	3	1	4	2
Expert	4	3	2	1
	GA138_Sig	**Gle_Col**	**NDH_Lav**	**NDH_Law**
Crash	3	4	2	1
Expert	2	4	1	3

Table E.1 (continued).

	GA138_Sig	Law_Suw	NDH_Lav	Sco_Cla
Crash	3	4	1	2
Expert	2	3	1	4
	GA10_Oak	**GA20_Wil**	**NDH_Lav**	**Sug_Buf**
Crash	3	2	1	4
Expert	3	2	1	4
	GA20_Wil	**Gle_Col**	**MLK_Bro**	**NDH_Lav**
Crash	3	2	4	1
Expert	2	3	4	1
	Cob_Gre	**GA10_Oak**	**Gra_Sce**	**Sug_Buf**
Crash	1	4	2	3
Expert	4	2	1	3
	Cob_Gre	**GA10_Oak**	**NDH_Law**	**Sug_Buf**
Crash	2	3	1	4
Expert	4	2	1	3

	Gra_Sce	NDH_Law	Ros_Wie	Sug_Buf
Crash	2	3	1	4
Expert	2	3	1	4
	Cob_Gre	**GA10_Oak**	**Sco_Cla**	**Sug_Buf**
Crash	1	3	2	4
Expert	4	2	1	3
	GA10_Oak	**Mem_Cov**	**Pon_Mor**	**Ros_Wie**
Crash	3	4	2	1
Expert	4	3	2	1
	GA138_Sig	**MLK_Bro**	**NDH_Lav**	**Ros_Wie**
Crash	3	1	2	4
Expert	3	4	1	2
	Nor_Tec	**Whi_Lin**	**Ros_Wie**	**Gra_Sce**
Crash	4	3	1	2
Expert	4	3	1	2

REFERENCES

1. AASHTO (2004). “A Policy on Geometric Design of Highways and Streets.” 5th Edition.
2. Aljanahi, A. A. M., Rhodes, A. H., and Metcalfe, A. V. (1999). ‘‘Speed, speed limits, and road traffic accidents under free flow conditions.’’ Accident Analysis and Prevention, Vol. 31, pp. 161–168.
3. Allen, B. L., Shin, B. T., and Cooper, P. J. (1978). “Analysis of Traffic Conflict Collisions.” Transportation Research Record 667, TRB, National Research Council, Washington D.C., pp. 67-74.
4. Arem, B. van, A. P. de Vos, and M.J.W.A. Vanderschuren. (1997). “The microscopic traffic simulation model MIXIC 1.3.” TNO Inro, Department of Traffic and Transport, Delft, the Netherlands, January 1997.
5. Archer, J., and Young, W. (2009). “Signal treatments to reduce heavy vehicle crash-risk at metropolitan highway intersections.” Accident Analysis and Prevention, Vol. 41, No. 3, pp. 404-411.
6. Atev, S., Arumugam, H., Masoud, O., Janardan, R., and Papanikolopoulos, N. P. (2005). “A vision-based approach to collision prediction at traffic intersections.” IEEE Transactions on Intelligent Transportation Systems, 6(4):416– 423.
7. Autey, J. (2012). “Before and after traffic safety evaluations using computer vision techniques.” Masters’ Thesis, University of British Columbia.

8. Baker, W. T., and Glauz, W. D. (1977). “The Traffic Conflicts Experience in the United States.” In Proceedings, First Workshop on Traffic Conflicts, Oslo, Norway.
9. Bauer, K.M., and Harwood, D. (1996). “Statistical Models of At-Grade Intersection Accidents.” Report No. FHWA-RD-96-125, McLean, Va., 1996.
10. Beymer, D., McLauchlan, P., Coifman, B., and Malik, J. (1997). “A real-time computer vision system for measuring traffic parameters.” Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR ’97), pages 495–501, Washington, D.C., USA. IEEE Computer Society.
11. Boonsiripant, S., Hunter, M., and Rodgers, R. (2011). “Speed Profile Variation as a Road Network Screening Tool.” Transportation Research Record: Journal of the Transportation Research Board; 2236, pp. 83-91.
12. Caliendo, C. (2007). “A crash-prediction model for multilane roads.” Accident Analysis and Prevention, Vol. 39(4), pp. 657-70.
13. Campbell, K., Joksch, H.C., and Green, P.E. (1996). “A bridging analysis for estimating the benefits of active safety technologies.” UMTRI-96-18 Final Report. University of Michigan, Transportation Research Institute. Ann Arbor, MI.
14. Campbell, K. (2008). “Collision Surrogates.” First Human Factors Symposium: Naturalistic Driving Methods and Analysis, Virginia Tech Transportation Institute.
15. CARE 9 User Manual, Center for Advanced Public Safety, University of Alabama, 2009.
16. Chatterjee, A., Everett, J. D., Reiff, B., Schwetz, T. B., Seaver, W. L., and Wegmann, F. J. (2003). “Tools for Assessing Safety Impact of Long-Range Transportation Plans

in Urban Areas." Center for Transportation Research, The University of Tennessee, Knoxville, TN.

17. Cheng, W., and Washington, S. (2005). "Experimental evaluation of hotspot identification methods." Accident Analysis and Prevention, Vol. 37, pp. 870–881.
18. Cheng, W., and Washington, S. (2008). "New Criteria for Evaluating Methods of Identifying Hot Spots." Transportation Research Record, 2083, pp. 76–85.
19. Chin, H. C., Quek, S. T., and Cheu, R. L. (1992). "Quantitative Examination of Traffic Conflicts." Transportation Research Record 1376, pp. 67-74.
20. Chin, H. C., and Quek, S. T. (1997). "Measurement of Traffic Conflicts." Safety Science, Vol. 26, No. 3, pp. 169-185.
21. Chin, H.C. and Quddus, M.A. (2003), "Modeling count data with excess zeros". Sociological Methods and Research, Vol. 32(1), 90-116.
22. Cooper, D., and Ferguson, N. (1976). "A Conflict Simulation Model." Traffic Engineering and Control, Vol. 17, pp. 306-309.
23. David, N., and Norman, J.R. (1975). "Motor Vehicle Accidents in Relation to Geometric and Traffic Features of Highway Intersection," Volume II, FHWA-RD-76-129, Federal Highway Administration, Washington, D.C.
24. Debnath, A. K., and Chin, H. C. (2010). "Navigational traffic conflict technique: a proactive approach to quantitative measurement of collision risks in port waters." Journal of Navigation, 63(1), pp. 137-152.
25. Deacon, J.A., C. V. Zegeer, and R.C. Deen. (1975). "Identification of Hazardous Rural Highway Locations." Transportation Research Record 543, pp. 16-33.

26. Dean, C., Lawless, J. F., and Willmot, G. E. (1989). “A mixed Poisson-inverse-Gaussian regression model.” Canadian Journal of Statistics.

27. Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., Perez, M. A., Hankey, J., Ramsey, D., Gupta, S., Bucher, C., Doerzaph, Z. R., Jermeland, J., and Knipling, R.R. (2006). “The 100-Car Naturalistic Driving Study: Phase II – Results of the 100-Car Field Experiment.” Report No. DOT-HS-810-593, National Highway Traffic Safety Administration.

28. Dixit, V. V., Pande, A., Abdel-Aty, M., and Radwan, E. (2011). “Quality of traffic flow on urban arterial streets and its relationship with safety.” Accident Analysis and Prevention, Vol. 43(5), 1610-6.

29. El-Basyouny, K., and Sayed, T. (2006). “Comparison of Two Negative Binomial Regression Techniques in Developing Accident Prediction Models.” Transportation Research Record, 1950, pp. 9–16.

30. Elvik, R. (2007). “State-of-the-Art Approaches to Road Accident Black Spot Management and Safety Analysis of Road Networks.” Report 883. Institute of Transport Economics, Oslo.

31. Elvik, R. (2008). “Comparative Analysis of Techniques for Identifying Hazardous Road Locations.” Transportation Research Record, 2083, pp. 72–75.

32. Elvik, R., Erke, A., and Christensen, P. (2009). “Elementary Units of Exposure.” Transportation Research Record, No. 2103, pp. 25-31.

33. Fancher, P., Ervin, R., Sayer, J., Hagan, M., Bogard, S., Bareket, Z., Mefford, M., Haugen,J. (1998). “Intelligent cruise control field operational test (Technical report

DOTHS 808 849).” National Highway Traffic Safety Administration, Washington, DC.

34. Fridstrøm, L., Ifver, J., Ingebrigsten, S., Kulmala, R., and Thomsen L. K. (1995). "Measuring the Contribution of Randomness, Exposure, Weather, and Daylight to the Variation in the Road Accident Counts." Accident Analysis and Prevention, Vol. 27(1), pp. 1-20.
35. Garber, N.J. and R. Gadiraju. (1989). “Factors affecting speed variance and its influence on accidents.” Transportation Research Record. Vol.1213, pp. 64-71.
36. Gettman, D., and Head, L. (2003). “Surrogate Safety Measures from Traffic Simulation Models.” Final Report, Federal Highway Administration.
37. Gettman, D., Pu, L., Sayed, T., and Shelby, T. (2008). “Surrogate Safety Assessment Model and Validation: Final Report.” Publication FHWA-HRT-08-051. FHWA, U.S. Department of Transportation.
38. Glauz, W. D., and Migletz, D. J. (1980). “Application of Traffic Conflict Analysis at Intersections.” NCHRP Report 219, Transportation Research Board, National Research Council, Washington D. C.
39. Guo, F., S. G. Klauer, J. M. Hankey, and T. A. Dingus (2010). “Using Near-Crashes as a Crash Surrogate for Naturalistic Driving Studies.” Transportation Research Record, Vol 2147, pp 66-74.
40. Haddon, W. (1972) “A Logical Framework for Categorizing Highway Safety Phenomena and Activity.” Journal of Trauma, Vol. 12, No. 3.

41. Hakkert, A. S., and Mahalel, D. (1978). “Estimating the number of accidents at intersections from a knowledge of the traffic flow on the approaches.” Accident Analysis and Prevention, Vol. 10, pp. 69-78.

42. Hauer, E. (1982). “Traffic Conflicts and Exposure.” Accident Analysis and Prevention. Vol. 14(5), pp. 359-364.

43. Hauer, E. (1983). “Reflections on methods of statistical inference in research on the effect of safety countermeasures.” Accident Analysis and Prevention, Vol. 15 (4), pp. 275–285.

44. Hauer, E., and Garder, P. (1986). “Research into the Validity of the Traffic Conflict Technique.” Accident Analysis and Prevention, Vol. 18, pp. 164-174.

45. Hauer, E. (2004). “Statistical road safety modeling.” Transportation Research Record: Journal of the Transportation Research Board. Vol.1897, pp. 81-87.

46. Hayward, J. (1972) “Near Miss Determination through Use of a Scale of Danger.” Report TTSC-7115. The Pennsylvania State University, University Park.

47. Herman, R. (1959). “Traffic Dynamics: Analysis of Stability in Car Following.” Operations Research. Vol.7, pp. 86.

48. Hogema, J. H., and Janssen, W. H. (1996). “Effects of intelligent cruise control on driving behaviour: a simulator study.” TNO Human Factors, Soesterberg, The Netherlands. Report TM-1996-C-12.

49. Huang, F., Liu, P., Yu, H., and Wang, W. (2012). “Identifying if VISSIM simulation model and SSAM provide reasonable estimates for field measured traffic conflicts at signalized intersections.” Accident Analysis and Prevention, Vol. 50, pp. 1014-24.

50. Hyde, T., and Wright, C. C. (1986). “Extreme value methods for estimating road traffic capacity.” Transportation Research Part B, Vol. 20, pp. 125-138.

51. Hyden, C. (1977). “A Traffic-Conflicts Technique for Examining Urban Intersection Problems.” In proceedings, First workshop on traffic conflicts, Oslo, Norway.

52. Hyden, C. (1987). “The development of a method for traffic safety evaluation: The Swedish Traffic Conflicts Technique” PhD thesis, Lund University of Technology, Lund, Sweden, Bulletin 70.

53. Hyden, C. (1996). “Traffic Conflicts Technique: State-of-the-art. Traffic Safety Work with Video Processing.” H. H. Topp. Kaiserslauten, Germany, University Kaiserslautern. Transportation Department.

54. Ismail, K. (2010). “Application of computer vision techniques for automated road safety analysis and traffic data collection.” PhD Thesis. Dept. of Civil Engineering, The University of British Columbia.

55. Jones, T.R. and Potts, R. B. (1962). “The Measurement of Acceleration Noise - A Traffic Parameter.” Operations Research. Vol.10, pp. 745-763.

56. Kanhere, N. K., Birchfield, S. T., Sarasua, W. A., and Whitney, T. C. (2006). “Real-Time Detection and Tracking of Vehicle Base Fronts for Measuring Traffic Counts and Speeds on Highways.” Transportation Research Record 1993, pp. 155-164.

57. Karim, E., and Tarek, S. (2009). “Collision prediction models using multivariate Poisson-lognormal regression.” Accident Analysis and Prevention, Vol. 41 (4), pp. 820-828.

58. Khan, S., Shanmugam, R., and Hoeschen, B. (1999). “Injury, Fatal, and Property Damage Accident Models for Highway Corridors.” Transportation Research Record: Journal of the Transportation Research Board, No. 1665, Washington DC, 84-92.

59. Kiefer, R. J., LeBlanc, D. J., and Flannagan, C. A. (2004). “Developing an inverse time-to-collision crash alert timing approach based on drivers' last-second braking and steering judgments.” Accident Analysis & Prevention 37(2): 295-303.

60. Kloeden, C. N., McLean, A. J., Moore, V. M., and Ponte, G. (1997). “Traveling Speed and the Risk of Crash Involvement.” NHMRC Road Accident Research Unit, The University of Adelaide.

61. Kockelman, K.M. and W.J. Murray. (2007). “Freeway speeds and speed variations preceding crashes, within and across lanes.” Journal of the Transportation Research Forum, pp. 1-19.

62. Konduri, S., and Sinha, K. C. (2002). “Statistical Models for Prediction of Freeway Incidents.” Proceedings of the Seventh International Conference on Applications of Advances Technology in Transportation Engineering, Cambridge, MA, 167-174.

63. Kruysse, H. (1991) “The Subjective Evaluation of Traffic Conflicts Based on an Internal Concept of Dangerousness.” Accident Analysis and Prevention, Vol. 23, No. 1.

64. Laureshyn, A., Svensson, Å., Hydén, C. (2010). “Evaluation of traffic safety based on micro-level behavioural data: Theoretical framework and first implementation.” Accident Analysis & Prevention, Vol. 42, No. 16, pp. 37–46.

65. Lave, C.A. (1985). “Speeding, Coordination, and the 55 MPH Limit.” In The American Economic Review. Vol.75, pp. 1159-1164.

66. Leong, W.H.J. (1973). "Relationship Between Accidents and Traffic Volumes at Urban Intersections." Journal of Australian Road Research Board, 5(3): 72-90.

67. Ma, J., Kockelman, K., and Damien, P. (2008). “A multivariate Poisson-lognormal regression model for prediction of crash counts by severity, using Bayesian methods.” Accident Analysis and Prevention, Vol. 40(3), pp. 964-75.

68. Maurin, B., Masoud, O., and Papanikolopoulos, N. P. (2005). “Tracking all traffic: computer Meyer, M. D., and Cambridge Systematics (2008). “Crashes vs. Congestion – What’s the Cost to Society?” Report.

69. McDonald, J. W. (1953). “Relation Between Number of Accidents and Traffic Volume at Divided-Highway Intersections.” Highway Research Board Bulletin 74, pp. 7-17.

70. “Vision algorithms for monitoring vehicles, individuals, and crowds.” Robotics & Automation Magazine, IEEE, 12(1):29–36.

71. Messelodi, S., and Modena, C. M. (2005), “A computer vision system for traffic accident risk measurement.” Advances in Transportation Studies, Vol. 7(B), pp. 51-66.

72. Minderhoud, M. M., and Bovy, P. H. L. (2001). “Extended Time-to-Collision Measures for Road Traffic Safety Assessment.” Accident Analysis and Prevention, Vol. 33, No. 1, pp. 89-97.

73. Montella, A. (2009). “A Comparative Analysis of Hot-Spot Identification Methods.” Accident Analysis and Prevention, Vol. 42(2), pp. 571-81.

74. “Next Generation SIMulation Fact Sheet.” (2004). Report FHWA-HRT-06-135.

75. Nelder, J. A., and Wedderburn, R. W. M. (1972). “Generalized Linear Models." Journal of the Royal Statistical Society A, Vol. 135, pp. 370-384.

76. Nicholson, A. J. (1985). "The variability of accident counts." Accident Analysis and Prevention, Vol. 17, No. 1, pp. 47-56.

77. Olson, P. L., Farber, G. (2003). "Forensic Aspects of Driver Perception and Response." Second Edition.

78. Parker, Jr. M.R., and Zegeer, C.V. (1989). "Traffic Conflict Techniques for Safety and Operations." Engineers Guide. Report FHWA-IP-026. FHWA, U.S. Department of Transportation.

79. Peesapati, L., Hunter, M. P., Rodgers, M. O., and Guin, A. (2011) "A Profiling Based Approach to Surrogate Safety Data Collection." In Proceedings of the 3rd International Conference on Road Safety and Simulation, Indianapolis.

80. Perkins, S., and Harris, J. (1967). "Criteria for Traffic Conflict Characteristics", Report GMR 632. General Motors Corporation.

81. Persaud, B., Lyon, C., and Nguyen, T. (1999). "Empirical Bayes Procedure for Ranking Sites for Safety Investigation by Potential for Improvement." Transportation Research Record, 1665, pp. 7–12.

82. Pickering, D., R.D. Hall, and M. Grimmer (1986). "Accidents at Rural T-Junctions." Research Report 65, Transport and Road Research Laboratory, United Kingdom, 1986.

83. Poch, M., and Mannering F. (1996). "Negative binomial analysis of intersection accident frequencies." Journal of Transportation Engineering, Vol. 122, No.2, 105-113.

84. Porter, B. E., Berry, T. D., and Harlow, J. (1999). "A Nationwide Survey of Red Light Running: Measuring Driver Behaviors for the 'Stop Red Light Running' Program." Report, Daimler Chrysler Corporation.

85. Resende, P.T.V., and Benekohal, R. F. (1997). "Development of Volume-to-Capacity Based Accident Prediction Models." Proceedings of Traffic Congestion and Traffic Safety in the 21st Century, Chicago, IL, 215-221.

86. Saunier, N., and Sayed, T. (2007). "Automatic road safety using video analysis." Transportation Research Record: Journal of the Transportation Research Board, 2019:57-64.

87. Saunier, N., and Sayed, T. (2008). "Probabilistic Framework for the Automated Analysis of the Exposure to Road Collision." Transportation Research Record, No. 2083, Washington D.C., pp. 96-104.

88. Sayed, T., and Zein, S. (1999). "Traffic conflict standards for intersections." Transportation Planning and Technology, 22:309-323.

89. Shankar V., Jovanis, P., Aguerde, J., and Gross, F. (2008). "Analysis of Naturalistic Driving Data: Prospective View on Methodological Paradigms." Transportation Research Record 2061, 1-9.

90. Shoarian-Sattari, K., and Powell, D. (1987). "Measured Vehicle Flow Parameters as Predictors in Road Traffic Accident Studies." Traffic Engineering and Control, Vol. 28, No. 6, pp. 328-335.

91. Smith, J.T., and McKenna, M.C. (2013). "A Comparison of Logistic Regression Pseudo R^2 Indices." Multiple Linear Regression Viewpoints, 2013, Vol. 39(2)

92. Solomon, D. R. (1964). "Accidents on Main Rural Highways Related to Speed, Driver and Vehicle." U.S. Department of Commerce, Federal Bureau of Highways, Washington D. C.

93. Songchitruksa, P., and Tarko, A. P. (2004). "Using Imaging Technology to Evaluate Highway Safety." Report FHWA/IN/JTRP-2004, Grant No. SPR-2663, Purdue University.

94. Songchitruksa, P., and Tarko, A. P. (2006). "The extreme value theory approach to safety estimation." Accident Analysis and Prevention, Vol. 38, No. 4, pp. 811-822.

95. St-Aubin, P., Miranda-Moreno, L., and Saunier, N. (2013). "An automated surrogate safety analysis at protected highway ramps using cross-sectional and before–after video data." Transportation Research Part C: Emerging Technologies, Volume 36, November 2013, Pages 284-295.

96. Summala, H. (1996). "Accident risk and driver behaviour." Safety Science, Vol. 22, No 1-3, pp. 103-117.

97. Svensson, A. (1998). "A method for analyzing the traffic process in a safety perspective." Doctoral Thesis, Lund University, Sweden.

98. Svensson, Å. and Hydén, C. (2006). "Estimating the severity of safety related behaviour." Accident Analysis & Prevention, Vol. 38, pp. 379-385.

99. Tarko, A., Davis. G., Saunier, N., Sayed, T., and Washington, S. (2009). "Surrogate Measures of Safety." White Paper.

100. Tarko, A. (2012). "Use of crash surrogates and exceedance statistics to estimate road safety." Accident Analysis and Prevention, Vol. 45, pp. 230-240.

101. Turner, S., and A. Nicholson. (1998). “Intersection Accident Estimation: The Role of Intersection Location and Non-Collision Flows.” Accident Analysis and Prevention, Vol. 30, No. 4, pp. 505-517.

102. “User Manual - Bluetooth Slipper for GPS Receiver HI-401BT”.

103. Vasquez, A. D. and Fraichard, T. (2004). “Motion prediction for moving objects: a statistical approach.” In Proc. of the IEEE Int. Conf. on Robotics and Automation, pages 3931–3936, New Orleans, LA (US).

104. Vogel, K. (2002). “A comparison of headway and time to collision as safety indicators.”

105. Vogt, A., and Bared, J. G. (1998). “Accident Models for Two-Lane Rural Roads: Segments and Intersections.” Report No. FHWA-RD-98-133, McLean, VA.

106. Webb, W. B. (1955). “The Relation Between Accidents and Traffic Volumes at Signalized Intersections.” Proceedings of Institute of Transportation Engineers, pp. 149-167.

107. Williams, M. J. (1981). “Validity of the traffic conflicts technique.” Accident Analysis and Prevention 13(2), pp. 133-135.

108. Zimmerman, G., Zimolong, B., and Erke, H. (1977). “The Development of Traffic Conflicts Technique in the Federal Republic of Germany.” In Proceedings, First Workshop on Traffic Conflicts, Oslo, Norway.

109. Zheng, Z., Ahn, S., and Monsere, C. M. (2010). “Impact of Traffic Oscillations on Freeway Crash Occurrences.” Accident analysis and prevention, Vol. 42, pp. 626-636.

110. Zegeer, C.V., and F.M. Council. (1992). “Safety Effectiveness of Highway Design Features, Vol. III: Cross Sections.” FHWA-RD-91-046. Federal Highway Administration.

111. Zegeer, C.V., and F.M. Council. (1994). “Safety Relationships Associated with Cross-Sectional Roadway Elements.” Transportation Research Record, No. 1512, pp. 39–36.

112. Zegeer, C.H., R.C. Deen, and J.G. Mayes. (1981). “Effect of Lane and Shoulder Widths on Accident Reduction on Rural Two-Lane Roads.” Transportation Research Record 806, pp. 33-43.